

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:32:24 ; Search time 212 Seconds
(without alignments)
2847.672 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMQNTCHRMSFPRGKCPKPR.....MREHREHQRBPDETDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980a.*
2: geneseqp1990a.*
3: geneseqp2000a.*
4: geneseqp2001a.*
5: geneseqp2002a.*
6: geneseqp2003a.*
7: geneseqp2003ba.*
8: geneseqp2004a.*
9: geneseqp2005a.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	8	ADQ96675 Human rib
2	7500	100.0	1374	8	ADQ96675 Human rib
3	7500	100.0	1374	9	ADQ96675 Human rib
4	7486	99.8	1374	8	ADQ17464 Human rib
5	7486	99.8	1374	8	ADQ17464 Human rib
6	7483.5	99.8	1373	7	ADQ17386 Human rib
7	6426	85.7	1200	9	ABU63361 Human rib
8	6062	80.8	1374	6	ABU63361 Human rib
9	4071	54.3	769	4	ABU63361 Human rib
10	2963	39.5	1327	4	ABU63361 Human rib
11	2847	38.0	541	6	ABU63361 Human rib
12	2519	33.6	486	6	ABU63361 Human rib
13	2463	32.8	466	8	ADQ96675 Human rib
14	1961	26.1	378	4	ADQ96675 Human rib
15	1961	26.1	378	4	ADQ96675 Human rib
16	1961	26.1	378	4	ADQ96675 Human rib
17	1961	26.1	378	4	ADQ96675 Human rib
18	1330	17.6	263	7	ADQ96675 Human rib
19	1330	17.6	263	7	ADQ96675 Human rib
20	1295.5	17.3	301	4	ADQ96675 Human rib
21	1289	17.2	257	4	ADQ96675 Human rib
22	1122	15.0	277	4	ADQ96675 Human rib
23	890	11.9	412	8	ADQ96675 Human rib
24	890	11.9	412	8	ADQ96675 Human rib

25	890	11.9	412	9	ADQ96675 Human rib	ADQ96675 Human rib
26	702	9.4	148	5	ADQ96675 Human rib	ADQ96675 Human rib
27	496	6.6	115	2	ADQ96675 Human rib	ADQ96675 Human rib
28	496	6.6	115	3	ADQ96675 Human rib	ADQ96675 Human rib
29	472	6.3	97	5	ADQ96675 Human rib	ADQ96675 Human rib
30	340	4.5	1151	4	ADQ96675 Human rib	ADQ96675 Human rib
31	312.5	4.2	241	7	ADQ96675 Human rib	ADQ96675 Human rib
32	311	4.1	560	4	ADQ96675 Human rib	ADQ96675 Human rib
33	309	4.1	266	9	ADQ96675 Human rib	ADQ96675 Human rib
34	307.5	4.1	229	5	ADQ96675 Human rib	ADQ96675 Human rib
35	303.5	4.0	406	4	ADQ96675 Human rib	ADQ96675 Human rib
36	301.5	4.0	1663	8	ADQ96675 Human rib	ADQ96675 Human rib
37	298.5	4.0	228	5	ADQ96675 Human rib	ADQ96675 Human rib
38	298.5	4.0	228	8	ADQ96675 Human rib	ADQ96675 Human rib
39	298.5	4.0	228	8	ADQ96675 Human rib	ADQ96675 Human rib
40	298.5	4.0	228	8	ADQ96675 Human rib	ADQ96675 Human rib
41	298.5	4.0	242	5	ADQ96675 Human rib	ADQ96675 Human rib
42	297.5	4.0	800	8	ADQ96675 Human rib	ADQ96675 Human rib
43	285.5	3.8	225	9	ADQ96675 Human rib	ADQ96675 Human rib
44	285.5	3.8	1527	8	ADQ96675 Human rib	ADQ96675 Human rib
45	285.5	3.8	1636	8	ADQ96675 Human rib	ADQ96675 Human rib

ALIGNMENTS

RESULT 1	ADQ96675	ADQ96675 standard; protein; 1374 AA.
XX	ADQ96675	
AC	ADQ96675	
XX	ADQ96675	
DT	23-SEP-2004	(first entry)
XX		
DB	Human ribonuclease III (RNase III) enzyme.	
XX		
KW	Human; ribonuclease III; RNase III; research purpose; biological purpose; clinical purpose; cellular interaction; enzyme.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..220
FT	Region	/note="Proline rich region"
FT	Misc-difference	18
FT	Region	221..470
FT	Region	/note="Encoded by CGT"
FT	Region	949..1374
FT	Domain	/note="Serine-arginine rich region"
FT	Misc-difference	1211
FT	Region	/note="RNAse III domain"
FT	Region	/note="Encoded by GCG"
FT	Region	1262..1269
FT	Region	/note="Alpha helix"
FT	Region	1282..1290
FT	Region	/note="Beta sheet"
FT	Region	1297..1303
FT	Region	/note="Beta sheet"
FT	Region	1308..1315
FT	Region	/note="Beta sheet"
FT	Region	1316..1336
FT	Region	/note="Alpha helix"
FT	Misc-difference	1345
FT	Misc-difference	/note="Encoded by GAA"
PN	US2004126867-A1.	
XX		
PD	01-JUL-2004.	
XX		
PF	09-FEB-2004; 2004US-00774974.	
XX		
PR	06-JUL-2001; 2001US-00900425.	
XX		

PA (CROO/) CROOKE S T.
 PA (WUH/) WU H.
 XX
 PI Crooke ST, Wu H;
 XX
 DR WPI, 2004-516913/49.
 XX N-PSDB; ADQ96674.
 XX
 PT New isolated nucleic acid molecule encoding human RNase III, useful for
 PT research, biological, or clinical purposes, e.g. defining the roles of
 PT RNase III and the interaction of human RNase III and cellular RNA.
 XX
 PS Claim 3; SEQ ID NO 2; 31pp; English.
 XX
 CC The invention relates to human ribonuclease III (RNase III) and its
 CC corresponding nucleic acid sequence. The polynucleotide sequence of the
 CC invention is useful for research, biological and clinical purposes. It is
 CC useful in defining the roles of RNase III and the interaction of human
 CC RNase III and cellular RNA. The present sequence is the human RNase III
 CC enzyme.
 CC
 SQ Sequence 1374 AA;
 Query Match 100.0%; Score 7500; DB 8; Length 1374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NMQGNTHMSFPHRGCGPRGCGHAGRPAPSPFQNTLRLHPQOPVQYQYEPSPAS 60
 1 MDAQGTCMHMSFPHRGCGPRGCGHAGRPAPSPFQNTLRLHPQOPVQYQYEPSPAS 60
 61 TTFSSNPAFNFLPRPDPVFPFPMPMPMPMPMPMPMPMPMPMPMPMPMPMPMPMP 120
 61 TTFSSNPAFNFLPRPDPVFPFPMPMPMPMPMPMPMPMPMPMPMPMPMPMPMPMP 120
 121 PRPMPCPNPVPVPGAPGCGTTPFPMPMPMPMPMPMPMPMPMPMPMPMPMPMPMP 180
 121 PRPMPCPNPVPVPGAPGCGTTPFPMPMPMPMPMPMPMPMPMPMPMPMPMPMPMP 180
 181 SFNSFQNNSSFPSPANNSSPHFRLLPYPPLPKASERRSPERLKGYDHRHRDHSGR 240
 181 SFNSFQNNSSFPSPANNSSPHFRLLPYPPLPKASERRSPERLKGYDHRHRDHSGR 240
 241 GERHSLDRRERGRSPDRRRODSRYSDYDRGRTSPRHSYERSRERERHRHRDRRS 300
 241 GERHSLDRRERGRSPDRRRODSRYSDYDRGRTSPRHSYERSRERERHRHRDRRS 300
 301 PSLERSYKKEYSRSGSYGLSVVPBPAGCTPELPGLIKNTDSWAPPLFIVNHRSPSRK 360
 301 PSLERSYKKEYSRSGSYGLSVVPBPAGCTPELPGLIKNTDSWAPPLFIVNHRSPSRK 360
 361 KQARWEEEDRMSDNOSSGDKNYTSIKEKEPEETMPDKNEEBEELKPMVIRCTHSEN 420
 361 KQARWEEEDRMSDNOSSGDKNYTSIKEKEPEETMPDKNEEBEELKPMVIRCTHSEN 420
 421 YYSSDMDVGVSTVVGTSRLRLDYDKFEBELGSRQEKAAARPMPPEPKTKLDEDLSS 480
 421 YYSSDMDVGVSTVVGTSRLRLDYDKFEBELGSRQEKAAARPMPPEPKTKLDEDLSS 480
 481 SSESCESDSDSTSSSSSDSEVFVIAEIKRKAHPRLDELMTYNDPGQMDGFLCKGA 540
 481 SSESCESDSDSTSSSSSDSEVFVIAEIKRKAHPRLDELMTYNDPGQMDGFLCKGA 540
 541 KARRTIRHSIYGEBAIKPCRPMTNAGRLFYRTITVSPPTFLDRPTVIEYDHEXT 600
 541 KARRTIRHSIYGEBAIKPCRPMTNAGRLFYRTITVSPPTFLDRPTVIEYDHEXT 600
 601 PEGFSMFAHAPLTNIPDLCKVIRENIDYTIHFIEEMMPENFCVAGLFSFLFRDLIELY 660
 601 PEGFSMFAHAPLTNIPDLCKVIRENIDYTIHFIEEMMPENFCVAGLFSFLFRDLIELY 660
 661 DNNLKGPLFEDSPCCPRPHFMRFPVRLPDGSEVLSHQIILYLIRCSKALVPEEETA 720
 661 DNNLKGPLFEDSPCCPRPHFMRFPVRLPDGSEVLSHQIILYLIRCSKALVPEEETA 720

DB 661 DNNLKGPLFEDSPCCPRPHFMRFPVRLPDGSEVLSHQIILYLIRCSKALVPEEETA 720
 QY 721 NMLQWEELEWQCYAECECKMTVTNPCTKSSVRIQDLDRQFNPDVTPTPIIVHGIRPA 780
 DB 721 NMLQWEELEWQCYAECECKMTVTNPCTKSSVRIQDLDRQFNPDVTPTPIIVHGIRPA 780
 QY 781 QLSYAGDPQY 840
 DB 781 QLSYAGDPQY 840
 QY 841 ELSSGQFMWTKTGRSDVCGHAWMLPVLTHIRHQCMLMDKLGITFORCLLQALMTNP 900
 DB 841 ELSSGQFMWTKTGRSDVCGHAWMLPVLTHIRHQCMLMDKLGITFORCLLQALMTNP 900
 QY 901 SHLNFQNMPPDHARNSLNSCGIRPKYGRKYNHNMKKGINTLINTMSRLGQDDPTPS 960
 DB 901 SHLNFQNMPPDHARNSLNSCGIRPKYGRKYNHNMKKGINTLINTMSRLGQDDPTPS 960
 QY 961 RINNHRELEFLGDANVEPLTSVHLVYLPSPSEGGATYRTAIYONQHLAMLAKKLELP 1020
 DB 961 RINNHRELEFLGDANVEPLTSVHLVYLPSPSEGGATYRTAIYONQHLAMLAKKLELP 1020
 QY 1021 FMLYAHGPDLCRESDLRMANCFEALIGAVYLEGSLBEAKOLFGRLLFNDDPLREVMIN 1080
 DB 1021 FMLYAHGPDLCRESDLRMANCFEALIGAVYLEGSLBEAKOLFGRLLFNDDPLREVMIN 1080
 QY 1081 YELHPLOQEPNTDROLIETSPVLQKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
 DB 1081 YELHPLOQEPNTDROLIETSPVLQKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTL 1140
 QY 1141 GHNQMEPLGDSIMOLVATEYLFIHPDHEHGLTLRSVLYNNRQAQVAEELGQOEYA 1200
 DB 1141 GHNQMEPLGDSIMOLVATEYLFIHPDHEHGLTLRSVLYNNRQAQVAEELGQOEYA 1200
 QY 1201 ITNDTKRPVGLRTYTLADLESFIALYTDKOLEVHTFMNVCFEPRLKEFLINQDMD 1260
 DB 1201 ITNDTKRPVGLRTYTLADLESFIALYTDKOLEVHTFMNVCFEPRLKEFLINQDMD 1260
 QY 1261 PKSQLOQCCLTLRTGKEBDIPLYKTLQTVGSHARTYVAVYFGERIGCGKSPISQQA 1320
 DB 1261 PKSQLOQCCLTLRTGKEBDIPLYKTLQTVGSHARTYVAVYFGERIGCGKSPISQQA 1320
 QY 1321 EMGAAMDALFKYFPQMAHQKFGIKRKYQELKEMRWEEHQBREDDEDIKK 1374
 DB 1321 EMGAAMDALFKYFPQMAHQKFGIKRKYQELKEMRWEEHQBREDDEDIKK 1374

RESULT 2
 ADR68727
 ID ADR68727 standard; protein; 1374 AA.
 XX
 AC ADR68727;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human Ribonuclease III, RNase III.
 XX
 KW Human; enzyme; Ribonuclease III; RNase III; RNA interference;
 KW gene silencing; double stranded RNA; pre-rRNA processing; RNA processing;
 KW RNA expression; RNA splicing; RNA translocation.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..220
 FT /label = Proline rich domain
 FT /note = "Claimed in claim 8"
 FT Domain 221..470
 FT /label = Serine Arginine rich domain
 FT /note = "Claimed in claim 8"
 FT Domain 499..1374
 FT /label = RNase III domain
 FT /note = "Claimed in claim 8"

FT Misc-difference 1211 /note= "Encoded by GCG"
 FT Misc-difference 1345 /note= "Encoded by GAA"
 FT
 XX US2004175828-A1.
 PD 09-SEP-2004.
 XX
 XX 22-MAR-2004; 2004US-00805919.
 XX
 XX 06-JUL-2001; 2001US-00900425.
 PR
 PA (CROO/) CROOKE S T.
 PA (WUHH/) WU H.
 PI Crooke ST, Wu H;
 XX
 XX WPI; 2004-667111/65.
 DR
 XX
 XX
 PT Modulating RNA interference, processing, expression, splicing and
 PT translocation in cell or tissue, by contacting cell or tissue with human
 PT RNAse III or oligomeric compound targeted to nucleic acid encoding human
 PT RNAse III.
 PS
 PS Claim 7; SEQ ID NO 2; 32pp; English.

CC The invention relates to modulating RNA interference (gene silencing),
 CC RNA processing (e.g. of pre-rRNA), RNA expression, RNA splicing, or RNA
 CC translocation in a cell or tissue, by contacting the cell or tissue with
 CC a modulator effective to cause the modulation by at least 50% as compared
 CC to control, where modulator is human RNAse III (ribonuclease III, which
 CC cleaves double stranded RNA) polypeptide or an oligomeric compound (an
 CC antisense oligonucleotide) targeted to the nucleic acid encoding human
 CC RNAse III. The oligomeric compound is targeted to a 3'-untranslated
 CC region (3'UTR), a 5'-untranslated region (5'UTR) or a coding region of a
 CC nucleic acid molecule encoding human RNAse III polypeptide, where the
 CC oligomeric compound inhibits the expression of human RNAse III
 CC polypeptide by at least 50%. In the method above, the RNA is rRNA, snRNA,
 CC snRNA or miRNA, or precursors of the above. The RNA is processed to form
 CC one or more 30S and 32S RNA fragments. The 32S RNA is processed to form
 CC one or more 12S pre-rRNA and 28S rRNA fragments. The methods are useful
 CC for modulating RNA interference in a cell or tissue, modulating
 CC processing of an RNA in a cell or tissue, modulating RNA expression in a
 CC cell or tissue, modulating RNA splicing in a cell or tissue, and
 CC modulating RNA translocation in a cell or tissue. The methods are useful
 CC for research, biological and clinical purposes. The methods are useful in
 CC defining the roles of RNAse III and interaction of human RNAse III and
 CC cellular RNA. The present sequence represents human RNAseIII.

XX Sequence 1374 AA;

Query Match 100.0%; Score 7500; DB 8; Length 1374;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMQAGTCHMSHPGCGPRGGGARGSPAPSPFRONRLAHPOQPVQYQYEPSPAS 60
 DB 1 MMQAGTCHMSHPGCGPRGGGARGSPAPSPFRONRLAHPOQPVQYQYEPSPAS 60
 QY 61 TTFSSNAPNPLPRPDPVFPFPPMPPSAQGLPPLPCPIRPPPNQMRHPFVPPCFPPM 120
 DB 61 TTFSSNAPNPLPRPDPVFPFPPMPPSAQGLPPLPCPIRPPPNQMRHPFVPPCFPPM 120
 QY 121 PPPMPCPNPVPYGAAPGQGTFFPMMPSPSMHPPPPPVMPQOVNQQYPPGYSHNFP 180
 DB 121 PPPMPCPNPVPYGAAPGQGTFFPMMPSPSMHPPPPPVMPQOVNQQYPPGYSHNFP 180
 QY 181 SENSNNSSFLPSANSSSPHRLPYPLPKASERSPERLKGHYDDHHRDHSR 240
 DB 181 SENSNNSSFLPSANSSSPHRLPYPLPKASERSPERLKGHYDDHHRDHSR 240
 QY 241 GERHRSLLDRERGRSPDRRRDORSYRSYDYGRTSPRHSYERSRERERHRHRDNRSS 300

DB 241 GERHRSLLDRERGRSPDRRRDORSYRSYDYGRTSPRHSYERSRERERHRHRDNRSS 300
 QY 301 PSLESYKKEYRSGSYGLSVPEPAGCTPELPGELIKNTQSMAPPLIIVNHRSPSRK 360
 DB 301 PSLESYKKEYRSGSYGLSVPEPAGCTPELPGELIKNTQSMAPPLIIVNHRSPSRK 360
 QY 361 KRARWEEKDRWSDNQSCKDNKNTSIKEKEPEETPDKNEEEEELKRWIRCTHSEN 420
 DB 361 KRARWEEKDRWSDNQSCKDNKNTSIKEKEPEETPDKNEEEEELKRWIRCTHSEN 420
 QY 421 YSSDPMDQVGSSTVYGTSLRLDLVDKFEELGSRQEKAKARPPMEPKTKLDEDESS 480
 DB 421 YSSDPMDQVGSSTVYGTSLRLDLVDKFEELGSRQEKAKARPPMEPKTKLDEDESS 480
 QY 481 SSECEDESDSTCSSSDSEVPDVAEIRKKAHPDRLDELMYNDPGQMDGPLCKCA 540
 DB 481 SSECEDESDSTCSSSDSEVPDVAEIRKKAHPDRLDELMYNDPGQMDGPLCKCA 540
 QY 541 KARRTGIRHSIYPGEBAIKPCRPMTNAGRLPHYRITVSPPTNFLTDRPTVIEYDHEX 600
 DB 541 KARRTGIRHSIYPGEBAIKPCRPMTNAGRLPHYRITVSPPTNFLTDRPTVIEYDHEX 600
 QY 601 FEGFSMFAHAPLTNIPLCVIRPNIDYTHFIEEMMPENFCYKGLSESLFPROLIELY 660
 DB 601 FEGFSMFAHAPLTNIPLCVIRPNIDYTHFIEEMMPENFCYKGLSESLFPROLIELY 660
 QY 661 DNMLGPLEDESPCCPRFHPMPFVPLPDGKEVLNHOGLLYLRCSKLVPEEEL 720
 DB 661 DNMLGPLEDESPCCPRFHPMPFVPLPDGKEVLNHOGLLYLRCSKLVPEEEL 720
 QY 721 NMLQWEELEWQYAECECKMIVTNPQTKPSVRIDLDREQNPDPVITPIIVFGIRPA 780
 DB 721 NMLQWEELEWQYAECECKMIVTNPQTKPSVRIDLDREQNPDPVITPIIVFGIRPA 780
 QY 781 QLSYAGDPQYQKLMSYVLRHLNLANSPVKQTDKOKLAQREBALQIKQKTMREEVTV 840
 DB 781 QLSYAGDPQYQKLMSYVLRHLNLANSPVKQTDKOKLAQREBALQIKQKTMREEVTV 840
 QY 841 ELSQGFMTQYGRSDVCOHAMMLPVLTNHIRVHOCIMHLDKIGYFQDRCLDOLAMTHP 900
 DB 841 ELSQGFMTQYGRSDVCOHAMMLPVLTNHIRVHOCIMHLDKIGYFQDRCLDOLAMTHP 900
 QY 901 SHHLNFGMNPDAHARNSLSNCGIROPKYGDRAKHHMMRKKGINTLINIMSRIGQDDPTS 960
 DB 901 SHHLNFGMNPDAHARNSLSNCGIROPKYGDRAKHHMMRKKGINTLINIMSRIGQDDPTS 960
 QY 961 RINNHERLEFLGDAVEFLTSVHLVYLPSPLEEGGIATYRTAIVONQHILAMLAKKLEDP 1020
 DB 961 RINNHERLEFLGDAVEFLTSVHLVYLPSPLEEGGIATYRTAIVONQHILAMLAKKLEDP 1020
 QY 1021 FMLYAHGPDLCRESOLRHNAMACFEALIGAVVLEGSLEBAKOLFGLLFPNDLDREVMIN 1080
 DB 1021 FMLYAHGPDLCRESOLRHNAMACFEALIGAVVLEGSLEBAKOLFGLLFPNDLDREVMIN 1080
 QY 1081 YPLHPLQLOEPRTDQLIETSPVLQKLTFFEEAIGVIFTHVBLARAFTLRVGFNHLTL 1140
 DB 1081 YPLHPLQLOEPRTDQLIETSPVLQKLTFFEEAIGVIFTHVBLARAFTLRVGFNHLTL 1140
 QY 1141 GHNQMEFLGDSIMQIVATEYLPFHPDHHGHLTLRSSLVNNRTQAKVAEELQMEYA 1200
 DB 1141 GHNQMEFLGDSIMQIVATEYLPFHPDHHGHLTLRSSLVNNRTQAKVAEELQMEYA 1200
 QY 1201 ITNDKTRPVGARTKTLADLESFALAYTDQLEEVHFMNVCFEPRLKEPFLINDQMD 1260
 DB 1201 ITNDKTRPVGARTKTLADLESFALAYTDQLEEVHFMNVCFEPRLKEPFLINDQMD 1260
 QY 1261 PKSQLOQCCLTARTGKEPDIPLYKTLQTVGSHARTVVAUYFGEIRIGCGKSPSIOQA 1320
 DB 1261 PKSQLOQCCLTARTGKEPDIPLYKTLQTVGSHARTVVAUYFGEIRIGCGKSPSIOQA 1320
 QY 1321 EYGAAMDALKXNPFOMAHQKRFIGRKYRQELKEMWEREHOEDBEDTEDIKK 1374

Db 1321 EMGAAMDALAEKYNPQMAHQKRFGRKRYKQBLKEMKMERERHQREBDEDEDIK 1374

RESULT 3
ID AEB47411 standard; protein; 1374 AA.
AC AEB47411;
XX 22-SEP-2005 (first entry)
DT Human ribonuclease III (RNase III) protein, SEQ ID NO: 2.
DE Ribonuclease; gene silencing; enzyme.
KM Homo sapiens.
OS
FH Key Location/Qualifiers
FT Region 1..220
FT /note= "RNase III proline-rich region"
FT Misc-difference 18
FT /note= "Encoded by CGT"
FT Region 221..470
FT /note= "RNase III serine-arginine rich region"
FT Domain 949..1374
FT /note= "RNase III domain"
FT Misc-difference 1211
FT /note= "Encoded by GCG"
FT Region 1262..1269
FT /note= "Alpha helix"
FT Region 1282..1290
FT /note= "Beta sheet"
FT Region 1297..1303
FT /note= "Beta sheet"
FT Region 1308..1315
FT /note= "Beta sheet"
FT Region 1318..1336
FT /note= "Alpha helix"
FT Misc-difference 1345
FT /note= "Alpha helix"
FT /note= "Encoded by GAA"
XX
XX US2005159384-A1.
XX
XX 21-JUL-2005.
XX
XX 02-DEC-2004; 2004US-00001993.
XX
XX 06-JUL-2001; 2001US-00900425.
XX PR 20-FEB-2002; 2002US-00079185.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Crooke ST;
XX
XX WPI; 2005-512270/52.
XX DR N-PSDB; AEB47410.
XX DR GENBANK; AAF80558.
XX
XX Use of RNase III for eliciting modification of a selected RNA target,
XX promoting gene silencing of a gene, inhibiting the expression of a gene,
XX promoting inhibition of expression of a gene, or eliciting modification
XX of an RNA target.
XX
XX Claim 19; SEQ ID NO 2; 32pp; English.
XX
XX The present invention relates to ribonuclease III (RNase III) nucleic
XX acids and their encoding proteins. RNase III is an endoribonuclease that
XX cleaves double stranded RNA. The invention is useful for eliciting
XX modification of a selected RNA target in a cell, promoting gene silencing
XX of a gene, inhibiting the expression of a gene, promoting inhibition of
XX expression of a gene and eliciting modification of a RNA target in a
XX cell. The present sequence is human ribonuclease III (RNase III) protein.

SQL Sequence 1374 AA:
Query Match 100.0%; Score 7500; DB 9; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMQNTGCHRMSPHRCGCGPRGGHGGARPSAPSPFNULRLHPQOPVQYQYEPSPSAS 60
DB 1 MMQNTGCHRMSPHRCGCGPRGGHGGARPSAPSPFNULRLHPQOPVQYQYEPSPSAS 60
QY 61 TTFSNSPAPNPLPPPPDDVPPPPPPPSAOGCLPPCPPIPPPPNQMRHPPPPCCFP 120
DB 61 TTFSNSPAPNPLPPPPDDVPPPPPPPSAOGCLPPCPPIPPPPNQMRHPPPPCCFP 120
QY 121 PPMPCPNPPVPVPGAPPGQGTFFPMMPSPSMHPPPPPPVPMQOVVYQYPPGYSHNFP 180
DB 121 PPMPCPNPPVPVPGAPPGQGTFFPMMPSPSMHPPPPPPVPMQOVVYQYPPGYSHNFP 180
QY 181 SFNSFQNNPSSFLPSANNSSPHFRLPPYLPKAPSERSEPERLKHVDHHRHDSHGR 240
DB 181 SFNSFQNNPSSFLPSANNSSPHFRLPPYLPKAPSERSEPERLKHVDHHRHDSHGR 240
QY 241 GERHSLDRBERGRSPDRRQDSRYRSDYDRGRTSPRRHSYRSRERERHRHRDRNS 300
DB 241 GERHSLDRBERGRSPDRRQDSRYRSDYDRGRTSPRRHSYRSRERERHRHRDRNS 300
QY 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPELPGEIINKTDSWAPLEIVNHRSPSRK 360
DB 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPELPGEIINKTDSWAPLEIVNHRSPSRK 360
QY 361 KPARWEEKDRMSDNQSSGKDNKNTYSIKKEPEEFMPDNNEEBELKPVNIRCHSN 420
DB 361 KPARWEEKDRMSDNQSSGKDNKNTYSIKKEPEEFMPDNNEEBELKPVNIRCHSN 420
QY 421 YSSDPMDQVGDSTVVGTSRLDLVDKFEEBELSGROEKAKAAPPMPEPKTLDLDELS 480
DB 421 YSSDPMDQVGDSTVVGTSRLDLVDKFEEBELSGROEKAKAAPPMPEPKTLDLDELS 480
QY 481 SESECESEDSTCSSSDSEVEFVLAERKKAHPDRDLDELWYNDPGMDGPLICSA 540
DB 481 SESECESEDSTCSSSDSEVEFVLAERKKAHPDRDLDELWYNDPGMDGPLICSA 540
QY 541 KARRTGIRHSTYEGEBAIKPCRPMTNNAGRLPHYATITSPPNPLTDRRTVLEYDHE 600
DB 541 KARRTGIRHSTYEGEBAIKPCRPMTNNAGRLPHYATITSPPNPLTDRRTVLEYDHE 600
QY 601 FEGFMPAHAPLTNTPLCKVIRFNIDYTHFIEEMMPENFCYKGELEFSLPRDILEY 660
DB 601 FEGFMPAHAPLTNTPLCKVIRFNIDYTHFIEEMMPENFCYKGELEFSLPRDILEY 660
QY 661 DNMLKGPLFEDSPCCPRFHPFRFVRPLPDGKEVLSMHQILLYLRCSSKALVPEE 720
DB 661 DNMLKGPLFEDSPCCPRFHPFRFVRPLPDGKEVLSMHQILLYLRCSSKALVPEE 720
QY 721 NMLQWBELEWQYAECEKMIYTNPGTKSSVRIOQLDREQNPDPVITPPIIVHFEIR 780
DB 721 NMLQWBELEWQYAECEKMIYTNPGTKSSVRIOQLDREQNPDPVITPPIIVHFEIR 780
QY 781 QLSYAGDPQYKLTMSYVYLRHLANSPKYKOTDKOKLQREBALOKIKQKTMRELEV 840
DB 781 QLSYAGDPQYKLTMSYVYLRHLANSPKYKOTDKOKLQREBALOKIKQKTMRELEV 840
QY 841 ELSQGFWMKTGIRSDVCGHMMPLVLTNHIRYHQCIMLHDLKIGYFQDRCLQLAMTH 900
DB 841 ELSQGFWMKTGIRSDVCGHMMPLVLTNHIRYHQCIMLHDLKIGYFQDRCLQLAMTH 900
QY 901 SHHLNFGANNPDAKNSLSNCGTRQKYGRKYNHMMKKGINTLINTSRIGQDDPTPS 960
DB 901 SHHLNFGANNPDAKNSLSNCGTRQKYGRKYNHMMKKGINTLINTSRIGQDDPTPS 960
QY 961 RINNHERLEFGDAVEFLTSVHLVYLPSSLREGGLATRTAYVONQHLAMLAKKLEDP 1020
DB 961 RINNHERLEFGDAVEFLTSVHLVYLPSSLREGGLATRTAYVONQHLAMLAKKLEDP 1020

QY 1021 FMVYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLBEAKQULFGRLTFNDPDLREVMIN 1080
DB 1021 FMVYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLBEAKQULFGRLTFNDPDLREVMIN 1080
QY 1081 YPLHPIQLOEPNTDROLITSPVLQKLTETEEBAIGVIFTHVRLLAFAFLTRVGFNHLTL 1140
DB 1081 YPLHPIQLOEPNTDROLITSPVLQKLTETEEBAIGVIFTHVRLLAFAFLTRVGFNHLTL 1140
QY 1141 GNNORHEFLGDSIMQVATEYLFTHFDDHBEGLTLRSLVNNRQAAVAEELGQOEYA 1200
DB 1141 GNNORHEFLGDSIMQVATEYLFTHFDDHBEGLTLRSLVNNRQAAVAEELGQOEYA 1200
QY 1201 INNDTKRPVGLRTKTLADLLESFIALYTDKOLEVHTFMNVCFFPRLKEFLINDMD 1260
DB 1201 INNDTKRPVGLRTKTLADLLESFIALYTDKOLEVHTFMNVCFFPRLKEFLINDMD 1260
QY 1261 PXSULOQCCLTLRTEGKEPDIPLYKTLQTVGSPSHARTYVAAYFKGERIGCGKPSIQQA 1320
DB 1261 PXSULOQCCLTLRTEGKEPDIPLYKTLQTVGSPSHARTYVAAYFKGERIGCGKPSIQQA 1320
QY 1321 EKGAMDALEKYNFPQMAHOKRIFGRKYQELKEMWREBHOEREDETEDIK 1374
DB 1321 EKGAMDALEKYNFPQMAHOKRIFGRKYQELKEMWREBHOEREDETEDIK 1374

RESULT 4

ADQ17464
ID ADQ17464 standard; protein; 1374 AA.

ADQ17464;

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated protein - SEQ ID 281.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.

XX Homo sapiens.

OS WO2004048938-A2.

PN 10-JUN-2004.

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Gineburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

DR WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression

PT of a gene in a first soft tissue sample and a normal soft tissue sample

PT and comparing the gene expression, also useful in treating soft tissue

PT sarcoma.

XX Example 2; SEQ ID NO 281; 210bp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma

CC which comprises obtaining a first soft tissue sample from an individual

CC and a normal soft tissue sample from the same or different individual,

CC determining the expression of a gene in both samples and comparing the

CC expression of the gene in both soft tissue samples, where a higher level

CC of protein expression in the first soft tissue sample indicates the

CC presence of soft tissue sarcoma. The method of the invention has

CC cytostatic applications and may be useful for detecting soft tissue

CC sarcoma, possibly via gene therapy and vaccine production. The nucleic

CC acid sequences may be useful in diagnostic and screening applications.

CC The current sequence is that of a human soft tissue sarcoma-upregulated

CC protein of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.

XX Sequence 1374 AA;

Query Match 99.8%; Score 7486; DB 8; Length 1374;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 1372; Conservative 0; Matches 2; Indels 0; Gaps 0;

QY 1 MMQAGTCHMSFHPGRCPRGRGAGAPSPAPFPQNRLLHPQAPVQYQYEPSAPS 60
DB 1 MMQAGTCHMSFHPGRCPRGRGAGAPSPAPFPQNRLLHPQAPVQYQYEPSAPS 60
QY 61 TTFSNSPAPNPLPRPDVPPPPMPBSAQGLPPCP1RPPFNHQRHPVPPPCPPM 120
DB 61 TTFSNSPAPNPLPRPDVPPPPMPBSAQGLPPCP1RPPFNHQRHPVPPPCPPM 120
QY 121 PPMPCPNPVPVPGAPPGQGTPPPMPPPSMHPPPPPVMPQVNYQYPPGYSHNPPPP 180
DB 121 PPMPCPNPVPVPGAPPGQGTPPPMPPPSMHPPPPPVMPQVNYQYPPGYSHNPPPP 180
QY 181 SFNSFQNNPSSFLPSANNSSSPHFRLLPYPLPKASERRSPRLKHYPDHRHDSHGR 240
DB 181 SFNSFQNNPSSFLPSANNSSSPHFRLLPYPLPKASERRSPRLKHYPDHRHDSHGR 240
QY 241 GERHRSIDRRERGRSPDRRRODSRYRSDYDRGTPSRHRSYERSRERERHRRHRRS 300
DB 241 GERHRSIDRRERGRSPDRRRODSRYRSDYDRGTPSRHRSYERSRERERHRRHRRS 300
QY 301 PSLERSYKKEKYGSGSGSYGLSVPEPAGCTPELPGI1KNTDSWAPPL1VNHRSRREK 360
DB 301 PSLERSYKKEKYGSGSGSYGLSVPEPAGCTPELPGI1KNTDSWAPPL1VNHRSRREK 360
QY 361 KARAMEEKDRMSDNOSSGDKNYTSIKKEPEETPDKNBEEBELKPMVIRCTHSEN 420
DB 361 KARAMEEKDRMSDNOSSGDKNYTSIKKEPEETPDKNBEEBELKPMVIRCTHSEN 420
QY 421 YSSDPMDOVGDSTVVGTSRLRDLYDKFEEELGSRQEKAAAPPEPKTLDLELSS 480
DB 421 YSSDPMDOVGDSTVVGTSRLRDLYDKFEEELGSRQEKAAAPPEPKTLDLELSS 480
QY 481 SSESCESDSDSTCSSSDSEVPDVI1AEIRKKAHPRLHDELMTYNDPGQNDGFLCKCSA 540
DB 481 SSESCESDSDSTCSSSDSEVPDVI1AEIRKKAHPRLHDELMTYNDPGQNDGFLCKCSA 540
QY 541 KARTGIRHSIYGEBAIKPCRPMTNNAAGL1FYRT1TVSP1N1FLDRPTV1EYDHEBY 600
DB 541 KARTGIRHSIYGEBAIKPCRPMTNNAAGL1FYRT1TVSP1N1FLDRPTV1EYDHEBY 600
QY 601 PEGFSMFAAPL1N1PLCKVIRFNIDYTH1FTEEMAPNFCVAGLE1FSL1FRD1LELY 660
DB 601 PEGFSMFAAPL1N1PLCKVIRFNIDYTH1FTEEMAPNFCVAGLE1FSL1FRD1LELY 660
QY 661 DNLKGP1FEDSPCCPRFHPMRFVRL1PDGKEVLSHGQ1LLYL1RCSKALVPEEEL1A 720
DB 661 DNLKGP1FEDSPCCPRFHPMRFVRL1PDGKEVLSHGQ1LLYL1RCSKALVPEEEL1A 720
QY 721 NMLQWELMOKYAEBCCKM1VTNPCTKSSVR1DOLDSEQFNPDVITP1IYHFGIRPA 780
DB 721 NMLQWELMOKYAEBCCKM1VTNPCTKSSVR1DOLDSEQFNPDVITP1IYHFGIRPA 780
QY 781 QLSYAGDPQYOKLMSYV1K1RHL1ANS1PKYQTDKOKL1QAREBAL1K1ROK1MREEV1Y 840
DB 781 QLSYAGDPQYOKLMSYV1K1RHL1ANS1PKYQTDKOKL1QAREBAL1K1ROK1MREEV1Y 840
QY 841 ELSSQGFMTGTGRSDVCQHMM1PVL1THIR1YHQC1MHL1DK1GYTFORCL1QLMTHP 900
DB 841 ELSSQGFMTGTGRSDVCQHMM1PVL1THIR1YHQC1MHL1DK1GYTFORCL1QLMTHP 900
QY 901 SHH1NGM1PDHARNS1SCG1ROPKYGRKYHMM1RKKG1INT1LN1MSRL1GQDDPTPS 960
DB 901 SHH1NGM1PDHARNS1SCG1ROPKYGRKYHMM1RKKG1INT1LN1MSRL1GQDDPTPS 960
QY 961 R1NHNERLEFLGDAVVEFLTSV1LYL1FSL1EEGLATYRTA1VONQ1HL1AKKLELDP 1020

|||||
 Db RINHERLEFLGDAVEFLTSVHLVLPSPLEGGIATRTAIYVONHILAMLAKKLELDP 1020
 QY 961 RINHERLEFLGDAVEFLTSVHLVLPSPLEGGIATRTAIYVONHILAMLAKKLELDP 1020
 QY 1021 EMLYHGPDLCESDLRHAMNANCEFLIGAVYLEGSEAKOLPGRLLFNDDLEWMLN 1080
 Db 1021 EMLYHGPDLCESDLRHAMNANCEFLIGAVYLEGSEAKOLPGRLLFNDDLEWMLN 1080
 QY 1081 YPLHPLQOEPMYTDRLIETSPVLQCLTEFEBAIGVIFTHVRLARAFTLRVGFNHLTL 1140
 Db 1081 YPLHPLQOEPMYTDRLIETSPVLQCLTEFEBAIGVIFTHVRLARAFTLRVGFNHLTL 1140
 QY 1141 GHNQMEFLGDSIMQVATEYLFIHPDHHEGHLTLIRSSLVNNRQAVASBELMQEYA 1200
 Db 1141 GHNQMEFLGDSIMQVATEYLFIHPDHHEGHLTLIRSSLVNNRQAVASBELMQEYA 1200
 QY 1201 IINDTKRVRGLRTKTLADLLESFIALATDQLEAVHFMNVCFPRLKEFLINDMDND 1260
 Db 1201 IINDTKRVRGLRTKTLADLLESFIALATDQLEAVHFMNVCFPRLKEFLINDMDND 1260
 QY 1261 PXSOLQCCCLTRTEGKEPDIPLYKTLQTVGSHARTYVAVYFKGERIGCGKGPISIOQA 1320
 Db 1261 PXSOLQCCCLTRTEGKEPDIPLYKTLQTVGSHARTYVAVYFKGERIGCGKGPISIOQA 1320
 QY 1321 EWGAAMDALCKNFPQMAHQKRFIGKTYQELKEMWEREHQREPEDETDIKK 1374
 Db 1321 EWGAAMDALCKNFPQMAHQKRFIGKTYQELKEMWEREHQREPEDETDIKK 1374
 QY 1374 EWGAAMDALCKNFPQMAHQKRFIGKTYQELKEMWEREHQREPEDETDIKK 1374
 Db 1374 EWGAAMDALCKNFPQMAHQKRFIGKTYQELKEMWEREHQREPEDETDIKK 1374

Sequence 1374 AA:
 Query Match 99.8%; Score 7486; DB 9; Length 1374;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MMQNTCHRMSPHPRGCGPRGGHGARGSPASFPFONILRLHPQOPVQYQYEPSPAS 60
 Db 1 MMQNTCHRMSPHPRGCGPRGGHGARGSPASFPFONILRLHPQOPVQYQYEPSPAS 60
 QY 61 TTFSSPAENFLPRPDPVPPPPMPPSAOGLPFPCPIRPPPPNQMRHPFVPPCFPM 120
 Db 61 TTFSSPAENFLPRPDPVPPPPMPPSAOGLPFPCPIRPPPPNQMRHPFVPPCFPM 120
 QY 121 PPMPCPNPPVPYGAAPGQGTFFPMMPSPMHPPPPPVPMQOVVQYQYPPGYSHANPPP 180
 Db 121 PPMPCPNPPVPYGAAPGQGTFFPMMPSPMHPPPPPVPMQOVVQYQYPPGYSHANPPP 180
 QY 181 SFNSFQNNPSSFLPSANNSSPHFNLPPYLPKAPSERSPERLKHVDHRRHDSHGR 240
 Db 181 SFNSFQNNPSSFLPSANNSSPHFNLPPYLPKAPSERSPERLKHVDHRRHDSHGR 240
 QY 241 GERHSLDRRERGRSDRRRDSRYRSDYDRGRTSPRRHSYRSRERERHRHRDRRS 300
 Db 241 GERHSLDRRERGRSDRRRDSRYRSDYDRGRTSPRRHSYRSRERERHRHRDRRS 300
 QY 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPELPGEIINKTDSMAPLEIVNHRSPSRBK 360
 Db 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPELPGEIINKTDSMAPLEIVNHRSPSRBK 360
 QY 361 KPARWEEKDRMSDNQSSGDKNYTSIKKEPEEFMPDNNEEBEELKPVNIRCHSEN 420
 Db 361 KPARWEEKDRMSDNQSSGDKNYTSIKKEPEEFMPDNNEEBEELKPVNIRCHSEN 420
 QY 421 YSSDPMDQVGDSTVGTSTRLDLYDKFEEELSGROEKAKAAPMEPEKTXLDELESS 480
 Db 421 YSSDPMDQVGDSTVGTSTRLDLYDKFEEELSGROEKAKAAPMEPEKTXLDELESS 480
 QY 481 SSECESDESDSTCSSSDSEVFDVIAEIKRKXHPDRILDELMYNDPQMGNDGPLCKCSA 540
 Db 481 SSECESDESDSTCSSSDSEVFDVIAEIKRKXHPDRILDELMYNDPQMGNDGPLCKCSA 540
 QY 541 KARRTGIRHSIYPGGEBAIKPCSPMTNNAGRLNHYRTIYSPFNPLTDRRTVIEYDHEYI 600
 Db 541 KARRTGIRHSIYPGGEBAIKPCSPMTNNAGRLNHYRTIYSPFNPLTDRRTVIEYDHEYI 600
 QY 601 FEGSPMPAARPLTNIPLCKVIRFNTDYTEHTFEBMMPENFCYKGLFPLFRDILEY 660
 Db 601 FEGSPMPAARPLTNIPLCKVIRFNTDYTEHTFEBMMPENFCYKGLFPLFRDILEY 660
 QY 661 DMNLKGPLFEDSPCCPRFHFMPFRVRLPDGGEVLSNHOILLVLRCSKALVPEEIA 720
 Db 661 DMNLKGPLFEDSPCCPRFHFMPFRVRLPDGGEVLSNHOILLVLRCSKALVPEEIA 720
 QY 721 NMLQWBELEWQYABCECKMITYNPGTKRSSVRIQDOLBREONPDVITPPIVHFGIRBA 780
 Db 721 NMLQWBELEWQYABCECKMITYNPGTKRSSVRIQDOLBREONPDVITPPIVHFGIRBA 780
 QY 781 OLSTAGDDPOYKLMKSYVYLRLHLANSPKVKOTDKQKLAQREBALOKIKQKTMRESEYV 840
 Db 781 OLSTAGDDPOYKLMKSYVYLRLHLANSPKVKOTDKQKLAQREBALOKIKQKTMRESEYV 840
 QY 841 ELSGSGFWTKGTRSDVCGHAMMLPYLTHIRYHOCIMLMDLIGYTFODRCILQIAMTHP 900
 Db 841 ELSGSGFWTKGTRSDVCGHAMMLPYLTHIRYHOCIMLMDLIGYTFODRCILQIAMTHP 900
 QY 901 SHHLNFGANPDHARNSLNCGRORPYGDKYVHNMHMKKGINTLINISRGQDDPTPS 960
 Db 901 SHHLNFGANPDHARNSLNCGRORPYGDKYVHNMHMKKGINTLINISRGQDDPTPS 960
 QY 961 RINHERLEFLGDAVEFLTSVHLVLPSPLEGGIATRTAIYVONHILAMLAKKLELDP 1020
 Db 961 RINHERLEFLGDAVEFLTSVHLVLPSPLEGGIATRTAIYVONHILAMLAKKLELDP 1020

QY 1021 FMUYAHGPDLCRESDDLRRHAMANCFEALIGAVYLEGSLIEBAKOLFGRLLFNDPDLREVMAN 1080
 DB 1021 FMUYAHGPDLCRESDDLRRHAMANCFEALIGAVYLEGSLIEBAKOLFGRLLFNDPDLREVMAN 1080
 QY 1081 YPLHPLQLOEPNTDROLIETSPVLOKLTETEEBAIGVIFTHVRLIADAFTLRITVGFMHLLT 1140
 DB 1081 YPLHPLQLOEPNTDROLIETSPVLOKLTETEEBAIGVIFTHVRLIADAFTLRITVGFMHLLT 1140
 QY 1141 GHNORHEFLGDSIMOLVATEYLFIHPPDHHEGLTLTSSLVNNRTOAKVAEELGMOEYA 1200
 DB 1141 GHNORHEFLGDSIMOLVATEYLFIHPPDHHEGLTLTSSLVNNRTOAKVAEELGMOEYA 1200
 QY 1201 ITNDTKRPEVGLRTKTLADLLESFIALYTDKLEVHTFPMVCFPPRLKEFLINDMND 1260
 DB 1201 ITNDTKRPEVGLRTKTLADLLESFIALYTDKLEVHTFPMVCFPPRLKEFLINDMND 1260
 QY 1261 PXSQLOCCLTLTETESKEPDIPLYKTLQTVGSPSHANTYVAVYFKGBRIGCGKPSIQQA 1320
 DB 1261 PXSQLOCCLTLTETESKEPDIPLYKTLQTVGSPSHANTYVAVYFKGBRIGCGKPSIQQA 1320
 QY 1321 EKGAAWDALAEKNFQMAHOKRFTGRKYRQELKEMNEREHOBERDEDEDIKK 1374
 DB 1321 EKGAAWDALAEKNFQMAHOKRFTGRKYRQELKEMNEREHOBERDEDEDIKK 1374
 RESULT 6
 ABU63361
 ID ABU63361 standard; protein; 1373 AA.
 AC ABU63361;
 XX 18-SEP-2003 (first entry)
 DT 18-SEP-2003 (first entry)
 DE Human double stranded RNase, RNase III.
 KM Human; enzyme; RNase III; double stranded RNase; RNA target;
 KW gene silencing.
 XX Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 18
 FT Misc-difference 1211 /note= "Encoded by CGT"
 FT Misc-difference 1345 /note= "Encoded by GCG"
 FT Misc-difference 1345 /note= "Encoded by GAA"
 PN US2003044941-A1.
 PD 06-MAR-2003.
 PF 20-FEB-2002; 2002US-00079185.
 PR 06-JUN-1996; 96US-00659440.
 PR 06-JUN-1997; 97US-00870608.
 PR 07-JAN-2000; 2000US-00479783.
 PR 06-JUL-2001; 2001US-00900425.
 PA (CROO/) CROOKE S T.
 PI Crooke ST;
 XX WPI; 2003-521756/49.
 DR N-PSDB; ACD27526.
 BLICiting a modification of a selected RNA target in a cell, useful for promoting inhibition of gene expression in a cell, comprises contacting an RNA-like polynucleotide-RNA target duplex with a polypeptide having an

XX The invention relates to eliciting a modification of a selected RNA target in a cell comprises contacting an RNA-like polynucleotide-RNA target duplex with a polypeptide comprising an RNase III domain. Also included are promoting gene silencing in a cell, inhibiting the expression of a gene in a cell comprising employing the method of cited above, promoting inhibition of expression of a gene, a hybrid RNase III (comprising at least one domain from a human RNase III and at least one domain from an RNase III of an organism other than human) and a cell having enhanced RNase III activity over an activity exhibited by a second cell (where the second cell is not enriched with respect to the amount or activity of RNase III polypeptide). The method is useful for eliciting a modification of a selected RNA target in a cell, and for promoting inhibition of expression of a gene in a cell. Compositions comprising RNase III polypeptides or polynucleotides are useful for research, biological and clinical purposes. The polynucleotides are may be used in defining the roles of RNase III and the interaction of human RNase III and cellular RNA. Host cells can be used for the production of human RNase III and for identifying agents, which increase or decrease levels of expression or activity of human RNase III in the cell. The present sequence represents human RNase III (a double stranded RNase)
 SQ Sequence 1373 AA;
 Query Match 99.8%; Score 7483.5; DB 7; Length 1373;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1373; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MNOGNTCHRMSPFPGGCGRGGGAGARPSABSFRRQNLRLHPQQPPVQYQIEPSPAPS 60
 DB 1 MNOGNTCHRMSPFPGGCGRGGGAGARPSABSFRRQNLRLHPQQPPVQYQIEPSPAPS 60
 QY 61 TTFNSPAPNFPLPRPDPVFPFPMPSPSAQPLPPCPPIRPPFNHOMRHPFPPCFPFM 120
 DB 61 TTFNSPAPNFPLPRPDPVFPFPMPSPSAQPLPPCPPIRPPFNHOMRHPFPPCFPFM 120
 QY 61 TTFNSPAPNFPLPRPDPVFPFPMPSPSAQPLPPCPPIRPPFNHOMRHPFPPCFPFM 120
 DB 61 TTFNSPAPNFPLPRPDPVFPFPMPSPSAQPLPPCPPIRPPFNHOMRHPFPPCFPFM 120
 QY 121 PPMPCPNPVPVGAAPGQGTFFPMMPSPSMRPPPPVMPQGVNYQYRPGYSHNFPPP 180
 DB 121 PPMPCPNPVPVGAAPGQGTFFPMMPSPSMRPPPPVMPQGVNYQYRPGYSHNFPPP 180
 QY 121 PPMPCPNPVPVGAAPGQGTFFPMMPSPSMRPPPPVMPQGVNYQYRPGYSHNFPPP 180
 DB 121 PPMPCPNPVPVGAAPGQGTFFPMMPSPSMRPPPPVMPQGVNYQYRPGYSHNFPPP 180
 QY 181 SFNSFONNPSSFLPSANNSSPFRLPPPLPKASERSPPRLKHYDHRDRDSHGR 240
 DB 181 SFNSFONNPSSFLPSANNSSPFRLPPPLPKASERSPPRLKHYDHRDRDSHGR 240
 QY 241 GERHSLDRERGRSPDRRQDSRYSDYDGRTPSRHSYERSREERERHHRDRRS 300
 DB 241 GERHSLDRERGRSPDRRQDSRYSDYDGRTPSRHSYERSREERERHHRDRRS 300
 QY 241 GERHSLDRERGRSPDRRQDSRYSDYDGRTPSRHSYERSREERERHHRDRRS 300
 DB 241 GERHSLDRERGRSPDRRQDSRYSDYDGRTPSRHSYERSREERERHHRDRRS 300
 QY 301 PSLERSYKKEYKRSYGLSVVPEBAGCTPBLPGELIKNTDSWAPPLEIVNHRSPSREK 360
 DB 301 PSLERSYKKEYKRSYGLSVVPEBAGCTPBLPGELIKNTDSWAPPLEIVNHRSPSREK 360
 QY 361 KRAREEEDKRWSDNOSGDKNYTISIKEKEPEETMPDKNEEBEELKVMIRCHSEN 420
 DB 361 KRAREEEDKRWSDNOSGDKNYTISIKEKEPEETMPDKNEEBEELKVMIRCHSEN 420
 QY 421 YSSDPMDOVGSSTVGTSTRDLVYKFEELGSRQEKAAAPRPEPPKTKLDEPLESS 480
 DB 421 YSSDPMDOVGSSTVGTSTRDLVYKFEELGSRQEKAAAPRPEPPKTKLDEPLESS 480
 QY 481 SESECEDESDTSCSSSDSEVPDVAIEIKRKAHPRLDELWYNDPGQNDGSLCKSA 540
 DB 481 SESECEDESDTSCSSSDSEVPDVAIEIKRKAHPRLDELWYNDPGQNDGSLCKSA 540
 QY 541 KARRTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTITVSPPTNPLTDRPVIYDYDHEYI 600
 DB 541 KARRTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTITVSPPTNPLTDRPVIYDYDHEYI 600
 QY 601 FSGFSMAHAPLNIPLCVIRINIDYTIHFIEMMPENPCVGLFSLFRDITIELY 660
 DB 601 FSGFSMAHAPLNIPLCVIRINIDYTIHFIEMMPENPCVGLFSLFRDITIELY 660
 QY 661 DWNHKGPLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRCSKALVPEBEIA 720

	661	721	721	721	781	781	781	841	841	901	901	961	961	1021	1021	1081	1081	1141	1141	1201	1201	1261	1261	1321	1321	1381	1381	1441	1441	1501	1501	1561	1561	1621	1621	1681	1681	1741	1741	1801	1801	1861	1861	1921	1921	1981	1981	2041	2041	2101	2101	2161	2161	2221	2221	2281	2281	2341	2341	2401	2401	2461	2461	2521	2521	2581	2581	2641	2641	2701	2701	2761	2761	2821	2821	2881	2881	2941	2941	3001	3001	3061	3061	3121	3121	3181	3181	3241	3241	3301	3301	3361	3361	3421	3421	3481	3481	3541	3541	3601	3601	3661	3661	3721	3721	3781	3781	3841	3841	3901	3901	3961	3961	4021	4021	4081	4081	4141	4141	4201	4201	4261	4261	4321	4321	4381	4381	4441	4441	4501	4501	4561	4561	4621	4621	4681	4681	4741	4741	4801	4801	4861	4861	4921	4921	4981	4981	5041	5041	5101	5101	5161	5161	5221	5221	5281	5281	5341	5341	5401	5401	5461	5461	5521	5521	5581	5581	5641	5641	5701	5701	5761	5761	5821	5821	5881	5881	5941	5941	6001	6001	6061	6061	6121	6121	6181	6181	6241	6241	6301	6301	6361	6361	6421	6421	6481	6481	6541	6541	6601	6601	6661	6661	6721	6721	6781	6781	6841	6841	6901	6901	6961	6961	7021	7021	7081	7081	7141	7141	7201	7201	7261	7261	7321	7321	7381	7381	7441	7441	7501	7501	7561	7561	7621	7621	7681	7681	7741	7741	7801	7801	7861	7861	7921	7921	7981	7981	8041	8041	8101	8101	8161	8161	8221	8221	8281	8281	8341	8341	8401	8401	8461	8461	8521	8521	8581	8581	8641	8641	8701	8701	8761	8761	8821	8821	8881	8881	8941	8941	9001	9001	9061	9061	9121	9121	9181	9181	9241	9241	9301	9301	9361	9361	9421	9421	9481	9481	9541	9541	9601	9601	9661	9661	9721	9721	9781	9781	9841	9841	9901	9901	9961	9961	10021	10021	10081	10081	10141	10141	10201	10201	10261	10261	10321	10321	10381	10381	10441	10441	10501	10501	10561	10561	10621	10621	10681	10681	10741	10741	10801	10801	10861	10861	10921	10921	10981	10981	11041	11041	11101	11101	11161	11161	11221	11221	11281	11281	11341	11341	11401	11401	11461	11461	11521	11521	11581	11581	11641	11641	11701	11701	11761	11761	11821	11821	11881	11881	11941	11941	12001	12001	12061	12061	12121	12121	12181	12181	12241	12241	12301	12301	12361	12361	12421	12421	12481	12481	12541	12541	12601	12601	12661	12661	12721	12721	12781	12781	12841	12841	12901	12901	12961	12961	13021	13021	13081	13081	13141	13141	13201	13201	13261	13261	13321	13321	13381	13381	13441	13441	13501	13501	13561	13561	13621	13621	13681	13681	13741	13741	13801	1380
--	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	------

[illegible]


```

Db      179 ----- 178
Qy      347 PLEIVNHSRPSREKKAARWEEKDRMSDNOSGKDXNTYSIKKBPETWPDKNBEEBE 406
Db      179 -----SRPSREKKAARWEEKDRMSDNOSGKDXNTYSIKKBPETWPDKNBEEBE 232
Qy      407 LKRWIRCTHSNNYSDDPMQVGSYVGTSLRLDLYDKFEELSGROEKAKAAPPM 466
Db      233 LKRWIRCTHSNNYSDDPMQVGSYVGTSLRLDLYDKFEELSGROEKAKAAPPM 232
Qy      467 EPPKTKLDDDLSSSSSECESEDSTCCSSSDSEVDVLAETKKAHPDRLDELMTND 526
Db      293 EPPKTKLDDDLSSSSSECESEDSTCCSSSDSEVDVLAETKKAHPDRLDELMTND 352
Qy      527 PGOMNDGPKCSAKARRTGIRHSIYPGEEAIKPCRPMTNNAARLPHYRTVSPNTPLT 586
Db      353 PGOMNDGPKCSAKARRTGIRHSIYPGEEAIKPCRPMTNNAARLPHYRTVSPNTPLT 412
Qy      587 DRPTVIEYDDHEXIFEGFSMFAHAPLTNIPLCVIRPNIDYTIHPIEMMPENFCVKGLE 646
Db      413 DRPTVIEYDDHEXIFEGFSMFAHAPLTNIPLCVIRPNIDYTIHPIEMMPENFCVKGLE 472
Qy      647 LPSLFLFRDILBYDNMLKGPLFEDSPCCPRFHFMRFRVRLPDGKKEVLSMHQILLYL 706
Db      473 LPSLFLFRDILBYDNMLKGPLFEDSPCCPRFHFMRFRVRLPDGKKEVLSMHQILLYL 532
Qy      707 LKCSKALVBEERANMLQWEELEWOKYABECKGMITNPGTKSSVRIDOLDEQGNPV 766
Db      533 LKCSKALVBEERANMLQWEELEWOKYABECKGMITNPGTKSSVRIDOLDEQGNPV 592
Qy      767 ITPPIIVHFGIRPAQSYAGDPQYQKLMKSYVGLRHLANSPKVKOTDKQKLAOREBALQ 826
Db      593 ITPPIIVHFGIRPAQSYAGDPQYQKLMKSYVGLRHLANSPKVKOTDKQKLAOREBALQ 652
Qy      827 KLRQKNTMRREVTVBELSOGFWKTGIRSDVCOHMMPLPVLTHIRYHOCMLDKLIGYT 886
Db      653 KLRQKNTMRREVTVBELSOGFWKTGIRSDVCOHMMPLPVLTHIRYHOCMLDKLIGYT 712
Qy      887 FODRCILOLAMTHPSHNLPGNMPDHARRSLSNCGIRQPKYGRKXHNHMRKKGINTLI 946
Db      713 FODRCILOLAMTHPSHNLPGNMPDHARRSLSNCGIRQPKYGRKXHNHMRKKGINTLI 772
Qy      947 NIMSRIGODDPTPSRINHNRELEFLGDAVVEFLTSVHLYYLPFSLBEGGLATRYTAIVON 1006
Db      773 NIMSRIGODDPTPSRINHNRELEFLGDAVVEFLTSVHLYYLPFSLBEGGLATRYTAIVON 832
Qy      1007 QHLAMLAKKLELDPFMLYAHGPDLCRESDLRHAMANCFEALIGAVYLGSLBBAKQLFGR 1066
Db      833 QHLAMLAKKLELDPFMLYAHGPDLCRESDLRHAMANCFEALIGAVYLGSLBBAKQLFGR 892
Qy      1067 LFPNDPDLAEWMLNLYLHLQLOEPNTDQLIETSVYLOKTEFEBAIGVIFPHVLLAR 1126
Db      893 LFPNDPDLAEWMLNLYLHLQLOEPNTDQLIETSVYLOKTEFEBAIGVIFPHVLLAR 952
Qy      1127 AFLTLRTVGFNHLTLGHNOBMEFLGDSIMQVATEYFIHPDHHBHLTLTLSSLVNNRT 1186
Db      953 AFLTLRTVGFNHLTLGHNOBMEFLGDSIMQVATEYFIHPDHHBHLTLTLSSLVNNRT 1012
Qy      1187 QAKVAEELQMOEYAITNDKTKRPVGLRTKTLADLBSFLAALYTDKOLEYVHTFMNVCF 1246
Db      1013 QAKVAEELQMOEYAITNDKTKRPVGLRTKTLADLBSFLAALYTDKOLEYVHTFMNVCF 1072
Qy      1247 PRLKSEILNQDNDPDSQLOCCCLTIRTEGKEDPILYKTLQVGSNARTTVAVYFPG 1306
Db      1073 PRLKSEILNQDNDPDSQLOCCCLTIRTEGKEDPILYKTLQVGSNARTTVAVYFPG 1132
Qy      1307 ERIGCGKSGISOAEKGAMDALEKYNFOMAHQKFKGRKROELKEMRWEREHOEREP 1366
Db      1133 ERIGCGKSGISOAEKGAMDALEKYNFOMAHQKFKGRKROELKEMRWEREHOEREP 1192
Qy      1367 DETEDIKK 1374

```

```

Db      1193 DETEDIKK 1200
RESULT 8
ABU08060
ID ABU08060 standard; protein; 1374 AA.
XX
AC ABU08060;
XX
DT 16-MAY-2003 (first entry)
XX
DE Human RNase III protein.
XX
KW Human; enzyme; RNase III; ribonuclease III; endoribonuclease;
KW pre-ribosomal RNA; pre-rRNA; small molecular weight nuclear RNA; snRNA;
KW small molecular weight nucleolar RNA; snorNA; mRNA degradation;
KW antisense therapy; RNA interference; RNAi; gene therapy;
KW infectious agent; prophylaxis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 151..168
FT Region /label= Alpha_helix_1
FT Region 181..189
FT Region /label= Beta_sheet_1
FT Region 196..202
FT Region /label= Beta_sheet_2
FT Region 207..214
FT Region /label= Beta_sheet_3
FT Region 215..235
FT Region /label= Alpha_helix_2
PN US2002164601-A1.
XX
PD 07-NOV-2002.
XX
PF 06-JUL-2001; 2001US-00900425.
XX
PR 06-JUN-1996; 96US-00659440.
PR 06-JUN-1997; 97US-00870608.
PR 07-JUN-2000; 2000US-00479783.
XX
PA (WUHH/) WU H.
PA (CROO/) CROOKE S T.
PI Wu H, Crooke ST;
XX
DR WPI; 2003-328390/31.
XX
PT New human RNase polypeptide, useful for screening antisense
PT oligonucleotides for therapy of disorders associated with RNase III
PT expression or activity, or for evaluating the efficacy of an antisense
PT therapy.
XX
PS Claim 3; Fig 1; 17bp; English.
XX
CC The invention discloses an isolated human ribonuclease III (RNase III)
CC polypeptide, and the nucleic acid encoding it. RNase III is an
CC endoribonuclease that cleaves double stranded RNA. A1 RNase III species
CC contain an RNase III signature sequence. RNase III has been reported to
CC be involved in the processing of pre-ribosomal RNA (pre-rRNA), small
CC molecular weight nuclear RNAs (snRNAs) and small molecular weight
CC nucleolar RNAs (snorRNAs), as well as the degradation of some mRNA
CC species. Also disclosed is an antibody targeted to the human RNase III
CC polypeptide, an antisense compound 8 - 50 nucleobases in length, which is
CC targeted to the nucleic acid encoding human RNase III polypeptide, and
CC methods for inhibiting human RNase III expression, or activity, in a cell
CC or tissue, identifying agents that increase or decrease the activity or
CC levels of the human RNase III polypeptide in a host cell, screening
CC oligonucleotides to identify effective antisense oligonucleotides for
CC inhibition of expression of a selected target protein, prognosticating
CC efficacy of antisense therapy of a selected disease, eliciting cleavage

```

3
CC of a selected cellular RNA target and promoting RNA interference (RNAi)
CC in a cell. The RNase III polypeptide, the polynucleotide encoding it and
CC the antisense oligonucleotides, are useful for gene therapy (e.g. for
CC treating a disease or disorder associated with RNase III expression or
CC activity, or associated with an infectious agent), prophylaxis or as
CC research reagents. The sequence presented is the human RNase III protein
CC
XX
Sequence 1374 AA:

Query Match 80.8%; Score 6062; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MMQNTCHMSFHPGRCGRGCGHAGPAPSPFRQNTLRLLHPQCPVQYQYEPPSAPS 60
DB 274 MMQNTCHMSFHPGRCGRGCGHAGPAPSPFRQNTLRLLHPQCPVQYQYEPPSAPS 333
QY 61 TTFSSPAPNPLPPRDPVPPPPPPPSAOGPLPCPIPPPPNMQMRHPPVPPCPFP 120
DB 334 TTFSSPAPNPLPPRDPVPPPPPPPSAOGPLPCPIPPPPNMQMRHPPVPPCPFP 393
QY 121 PPMPCPNPPPVPGAPPGCGTFPPMMPPSMHPPPPPVMPQOVNYPYPGYSHNFP 180
DB 394 PPMPCPNPPPVPGAPPGCGTFPPMMPPSMHPPPPPVMPQOVNYPYPGYSHNFP 453
QY 181 SFNSFQNNSSFLPSANSSSPHFRLLPPYPLPKASERSRSPRLKHVDHRHDSHGR 240
DB 454 SFNSFQNNSSFLPSANSSSPHFRLLPPYPLPKASERSRSPRLKHVDHRHDSHGR 513
QY 241 GERHSLDPRERGRSPRRQDSRYSDYDRGTRPFRHSYERSRREERHRHDSHGR 300
DB 514 GERHSLDPRERGRSPRRQDSRYSDYDRGTRPFRHSYERSRREERHRHDSHGR 573
QY 301 PLEERSYKKEYSKGRSGSYGLSVPEBAGCTPELPGRIIKNTDSWAPLEIVNHRSPREK 360
DB 574 PLEERSYKKEYSKGRSGSYGLSVPEBAGCTPELPGRIIKNTDSWAPLEIVNHRSPREK 633
QY 361 KARWEEKDRMSDNGSSGKDKNTYSIKKEPEETMPDKNEEBEELKPVWIRCTHSN 420
DB 634 KARWEEKDRMSDNGSSGKDKNTYSIKKEPEETMPDKNEEBEELKPVWIRCTHSN 693
QY 421 YSSSDPMDOVGVSTVGTSLRLDLVKFEELGSRQEKAKARPPEPEPKTLDLELSS 480
DB 694 YSSSDPMDOVGVSTVGTSLRLDLVKFEELGSRQEKAKARPPEPEPKTLDLELSS 753
QY 481 SSECESDSDTSCSSSSDSEVFVIAEIKRKAHPRLDELWYNDPGQNDGPLCKCSA 540
DB 754 SSECESDSDTSCSSSSDSEVFVIAEIKRKAHPRLDELWYNDPGQNDGPLCKCSA 813
QY 541 KARRTGIRHSIYPGEEBAIKPCRPMTNNAARLPHYRITVSPPTNFLTDRPTVIEYDHEXI 600
DB 814 KARRTGIRHSIYPGEEBAIKPCRPMTNNAARLPHYRITVSPPTNFLTDRPTVIEYDHEXI 873
QY 601 FEGFSWFAAPLPLNPLCKVIRNIDYTHIFLEEMPEMFVYKAGLELFLFRDLLEY 660
DB 874 FEGFSWFAAPLPLNPLCKVIRNIDYTHIFLEEMPEMFVYKAGLELFLFRDLLEY 933
QY 661 DNLIKPLFEDSPCCPRFHFMPFRVPLPDGKEVLSMHQILLYLRSKALVPEEBIA 720
DB 934 DNLIKPLFEDSPCCPRFHFMPFRVPLPDGKEVLSMHQILLYLRSKALVPEEBIA 993
QY 721 NMLQWEELEWQKAECECKMIVTNPCTKPSVRIIDLREQNPDVITPIIVHFGIRPA 780
DB 994 NMLQWEELEWQKAECECKMIVTNPCTKPSVRIIDLREQNPDVITPIIVHFGIRPA 1053
QY 781 QLSYAGDPOYQKLMKSYVYLRLHLANSPKYKQDKOKLAQREBALOKIKQKTMREVV 840
DB 1054 QLSYAGDPOYQKLMKSYVYLRLHLANSPKYKQDKOKLAQREBALOKIKQKTMREVV 1113
QY 841 ELSOGFWMTGTGRSDVCOHAMMLPVLTHHRIHYQCMLHLDKILIGYFODRCILLQMTMP 900
DB 1114 ELSOGFWMTGTGRSDVCOHAMMLPVLTHHRIHYQCMLHLDKILIGYFODRCILLQMTMP 1173
```

```
QY 901 SHHLNFGMNPDAARNSLNCGRKQKGRKHHNMRKKGINTLINIMSRIGODDPTPS 960
DB 1174 SHHLNFGMNPDAARNSLNCGRKQKGRKHHNMRKKGINTLINIMSRIGODDPTPS 1233
QY 961 RINHERLEFLGDVAVEEFLTSVHLTYLFPSSLEGGATATRTAIVONOLAMLAKEULD 1020
DB 1234 RINHERLEFLGDVAVEEFLTSVHLTYLFPSSLEGGATATRTAIVONOLAMLAKEULD 1293
QY 1021 FMLYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLERAKOLFGRLLFNDPDLREVMIN 1080
DB 1294 FMLYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLERAKOLFGRLLFNDPDLREVMIN 1353
QY 1081 YPLHPLQLOEPTDRLIETS 1101
DB 1354 YPLHPLQLOEPTDRLIETS 1374
```

```
RESULT 9
AAB92635
ID AAB92635 standard; protein; 769 AA.
XX
AC AAB92635;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:10949.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI, 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8: SEQ ID NO 10949; 2537bp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC complementary sequences defined in the specification, where the
CC complementary sequences defined in the specification, or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
```

CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX Sequence 769 AA;

Query Match 54.3%; Score 4071; DB 4; Length 769;
 Best Local Similarity 99.5%; Pred. No. 4.2e-280;
 Matches 765; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 606 MFAHAPLNTNPLCKVIRFNIDYTHIFIEEMMPNFCVKGLESLFLFRDILELYDMNLK 665
 DB 1 MFAHAPLNTNPLCKVIRFNIDYTHIFIEEMMPNFCVKGLESLFLFRDILELYDMNLK 60
 QY 666 GRLFEDSPCCPRFHFMPFRVRLPDGKEVLSHQQLLYLRLCSKALVEEBEIANMLQW 725
 DB 61 GRLFEDSPCCPRFHFMPFRVRLPDGKEVLSHQQLLYLRLCSKALVEEBEIANMLQW 120
 QY 726 EELMOKVAEECKMTVTNPTSPSVRIDQLDREOPNPVTFPIIVHGRIPAOLSYA 785
 DB 121 EELMOKVAEECKMTVTNPTSPSVRIDQLDREOPNPVTFPIIVHGRIPAOLSYA 180
 QY 786 GDEYOYKLMKSYKLBHLLANSPKVKQTDKQKLAQREALQKTRKNTMRREYTVELSSQ 845
 DB 181 GDEYOYKLMKSYKLBHLLANSPKVKQTDKQKLAQREALQKTRKNTMRREYTVELSSQ 240
 QY 846 GFAPKTSIRSDVCHAMMLPVLTHIRYHOCMLDKLIGTFODRCILLQAMTHPSHHLN 905
 DB 241 GFAPKTSIRSDVCHAMMLPVLTHIRYHOCMLDKLIGTFODRCILLQAMTHPSHHLN 300
 QY 906 FGNNPRAHRSLSNCGIRPKYGDGRKYNHNMKKGINTLINMSRLGODDPTPSRINHN 965
 DB 301 FGNNPRAHRSLSNCGIRPKYGDGRKYNHNMKKGINTLINMSRLGODDPTPSRINHN 360
 QY 966 ERLEFLGDAAVEFLTSVHLTYLLFPLSEEGSLATYRTAIVONQHILAMLAKELEDPFMLYA 1025
 DB 361 ERLEFLGDAAVEFLTSVHLTYLLFPLSEEGSLATYRTAIVONQHILAMLAKELEDPFMLYA 420
 QY 1026 HGPDLGRESDDLHMANCFEALIGAVYLGSLSEBAKQLFGRLLPNDPLREVMILNPLHP 1085
 DB 421 HGPDLGRESDDLHMANCFEALIGAVYLGSLSEBAKQLFGRLLPNDPLREVMILNPLHP 480
 QY 1086 LQIQENTDROQLIETSPVLOKLTPEFBAIGVIFTHRRLLARATTLRVGSHNLTLGNOR 1145
 DB 481 LQIQENTDROQLIETSPVLOKLTPEFBAIGVIFTHRRLLARATTLRVGSHNLTLGNOR 540
 QY 1146 MEPLGDSIMQVATEVYLFIHPDHEGHLTLRSSLVNNRTQAKVAEELMOEYATINDK 1205
 DB 541 MEPLGDSIMQVATEVYLFIHPDHEGHLTLRSSLVNNRTQAKVAEELMOEYATINDK 600
 QY 1206 TKRPVGLRTKTLADLESFIALYTDKLELYVTFMNVCFFPRLKEFILNQDNDPKSQL 1265
 DB 601 TKRPVGLRTKTLADLESFIALYTDKLELYVTFMNVCFFPRLKEFILNQDNDPKSQL 660
 QY 1266 QOCCLTLRTBEGKRPDIPLYKTLQTVGSHARTTVAVYFKGERIGCGKGSIOQAENGAA 1325
 DB 661 QOCCLTLRTBEGKRPDIPLYKTLQTVGSHARTTVAVYFKGERIGCGKGSIOQAENGAA 720
 QY 1326 MDALFKYNFQMAHOKRFIGRKYROELKEMRWEREHOEREPDETEDIKK 1374
 DB 721 MDALFKYNFQMAHOKRFIGRKYROELKEMRWEREHOEREPDETEDIKK 769

RESULT 10
 ID AAB58539 standard; protein; 1327 AA.
 AC AAB58539;
 XX
 DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 2409.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001MO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL02642.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 2409; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AAB57737-
 CC AAB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1327 AA;

Query Match 39.5%; Score 2963; DB 4; Length 1327;
 Best Local Similarity 45.6%; Pred. No. 4.7e-201;
 Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;

QY 90 QGFLPPCPILRPFPNHQMRHFPVPCFPMPMPFCPN-NRPVCGAP-----P 137
 DB 3 QGFLPPCPVQPA-----PPPPPPPEEDLSPGQGVPSHNSNSHSQ 46
 QY 138 GQGTFFPMRPPRPMRPPRPVPCQVNVQYPRGY-----SHNFRP 179
 DB 47 SSKSLDYVPEPAPYASS--VPSYDYPQOPAYEGEYVAYNEQAOKTGGQSHYQY 103
 QY 180 PSFNSPONNPSFLPANSSSPHRLPPYPL---PKASERARRERLKHVDHRRHD 235
 DB 104 PA-----SGSSTLYS-----YKPRDLYPSSNYPSPSRQ-----RYTS 139
 QY 236 HSHRGERRHSLDRRGRSPDRRQDSRYSDYDGRTPSRHRSYRSRERERERHRR 295
 DB 140 NSSGQYHNYP-----GYSSGR-----RYEQSHDQ-----EHRQIDSRAYAHBPBGHY 183
 QY 296 DNRSPSLRSYKKEYKRSRSGLSVPEPACCTELPGEILIKNTDMSAPLELYNHS 355
 DB 184 AHRQAQSGOHGY---YGAARN-----QVSDDSYRGHHERERN 219
 QY 356 PSREKGRK--RWEERKOR-----WSDNOSGKDKNTYSIKEKEPEETMPDNSEEBELLK 409
 DB 220 ETIEKTRAKRYVTEHDLRLQWCSNFC-----EKEDYVKKGNALSEADAV 267
 QY 410 PVMIRCTSHSENYSSDPMQOVGDSTVVGTSRLDLVDKPEEELGSGOEKAKAARPMWEPP 469

```
Db 268 ESWVRSSPAELYYERK-----SENEVRGARLQKCLTLPDELLQRAKVRKLPYVVP 323
Qy 470 KTKL-----DEDLSSSESECEDEDSTCSSSDSEVPD-----IAEIKKKAH 514
Db 324 PKRARRVCVKHKHSEACSSSSSSDDSDDEDA-----FKIDCCMEELSHKVOH 373
Qy 515 PDRLHELWYNDGQWNDGFLCKCSAKARTGIRHSIYRGEELIKRCRPTNNAGLFLY 574
Db 374 PQRVHADLWHDNDGEMNDGFLCRCSAKSRIGIRHIGIPEETGYKLCDPSSNAGLFLY 433
Qy 575 RIVSPPTNPLTPRVIEYEDHDEYIPEGSFMAHAPLTNIPLCYIRFNIDYIHIEE 634
Db 434 RISISPTNPLTPTPIIKHDEHEPLPEGSLSHRLSLPVCXIRIRIETIIEYEE 493
Qy 635 MPEPENCYKGLBELSFLPRDILELYDWMKGP-----PEDSPCCPRHFMRVFRFLP 690
Db 494 KQPENTIHELDIFPKYLFHELELVDFNLMPRLPGSNVBS---CPAFHFFRFRVDP 550
Qy 691 DGGKEVLSMHQILLYLRGSKALVPEBELANMIQMELEWQYAECKGMIVNPGTKS 750
Db 551 DNGKEVLAWEVLYRLLDNSAQLVERQQLHLNQISQSEWQNVVDIKGLVYKPGYKGC 610
Qy 751 SVRIDLDREQFNPV-----ITPPIYHGGIRPAQLSYAGDPQYOKLMSYVKLR 801
Db 611 SLRVQDLDRN--NSDLPECVDRETGISHPAIVHFGICHPOLSYAGNPEYOKAMREYVKTR 668
Qy 802 HLANSRKVKQTDQKLAQREBALOKIROKNTMRREVVELSSQGFMTGIRSDVOHAM 861
Db 669 HLMANNSKSPFQDKRLKEKEGRLQEMRQTKGRKKNITVAISSEGGYRGIQMDVVQHAM 728
Qy 862 MLPVLTNIRYHOCMLMDKLGITVRODRCLDLANTHSHHLNFGMNPDHANSLSNG 921
Db 729 LIPVLGHRFHKSLLDLSESIYGRKNRYLQDLALTHBKYKNYGTNPDHANSITNG 788
Qy 922 IQPKGDRKVNHMNRKKGINTLINIMSRIGQDDPTPSINNESEFLGDAVVEFLTS 981
Db 789 IQPEKGDRIKINMTRKGINLTVISMSEGENHETVSITNEHELEFGDAVVEFLSS 848
Qy 982 VHLVYLPFSLSEGGATYRTAIVQNOHMLAKKELDPRMLYAHGPDLCRESDLRHMA 1041
Db 849 IHLFMFPELEBEGGLATYRAIVQNOHMLAKKQLQLEEFMLYAHGSDLCHELELRHMA 908
Qy 1042 NCFEALIGAVYLEGSLEBAKOLFGRLLF--NDPDLREVMANYPHPIQLOEPNTDROLIET 1100
Db 909 NCFEALMGALLLDGGIKVADDEVFTDALFRDEKLLSIWKLPEHPLOEQEPRLGDRSCIDS 968
Qy 1101 SPVLQKLTPEFEBAIGVIFTHVRLARAFPLRTVGFPHNLTHGNQRMELGDSIMQVATE 1160
Db 969 YRVLKELTPEEDSIGIKFKHIRLARAFTDRSIGFTHLTIGSNQRLFEGLDVLQLCSE 1028
Qy 1161 YLFHFRPDHHEGHTLRLSSLVNNRTQAKVAEELGMEQVAI--TNDKTKRPVGLRTKTLAD 1219
Db 1029 YLVRHFRHNEHGLSLRLSSLVNNRTQAVVCDLGMKPAVYANPK-----ADLKTDRAD 1084
Qy 1220 LLESFIALYTDKDLBYVHTPMANVCFFRPRLKERTLIMQWMDPKSQIQCCCLTIRT--EGKE 1278
Db 1085 LLEAFIGALYVDKGLLYCEQFCVCLFPRLOLPIFMQWMDPKSKIQCCCLTIRTDGSE 1144
Qy 1279 PDIPLKTIQTVPSHARTYTVAVYFKGEVIGGKSPSTQQAAMGAAMALEKYN--PFO 1336
Db 1145 PDIPLYKVVASPTNTRYKVAVYFRSKLATSSGSSIQQAAMNAKQALENSRDLFPO 1204
Qy 1337 MAHQKRFIGKRYQEL---KEMWEREHOERREDETDIHK 1374
Db 1205 LDHQKRVIAKSIKKQTNELNDSDRQHDE-----EKIKR 1239
```

```
RESULT 11
ABM04841
ID ABM04841 standard; protein; 541 AA.
AC AC
XX ABM04841;
XX
```

```
DT 22-SEP-2003 (first entry)
XX Human putative ribonuclease III.
DE
XX
XX spinal cord; neuropathic pain; central sensitisation pain; pain;
KW analgesic; gene therapy.
XX
XX Homo sapiens.
XX
XX BP1284298-A2.
XX
XX 19-FEB-2003.
XX
XX 26-JUL-2002; 2002EP-00255229.
XX
XX 27-JUL-2001; 2001GB-00018354.
PR 07-FEB-2002; 2002GB-00002883.
XX
XX (WARN ) WARNER LAMBERT CO.
PA
XX Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
PI
XX WPI; 2003-543489/52.
DR
XX N-PSDB; ACF25385.
XX
XX Use of an isolated gene sequence in the screening of compounds for
PT diagnosing or treating pain.
PS Claim 1; Page 183-184; 188pp; English.
XX
XX The invention relates to a novel isolated gene sequence that is
CC downregulated in the spinal cord of a mammal in response to mechanically
CC distinct first and second models of neuropathic or central sensitisation
CC pain, useful in the screening of compounds for diagnosing or treating
CC activity. A protein encoded by a gene of the invention has analgesic
CC activity. The gene sequence is useful for preparing a composition for
CC diagnosing or treating pain. The present sequence represents a protein
CC encoded by a gene of the invention
CC
SQ Sequence 541 AA;
Qy 834 MREVTVELSSQGFMTGIRSDVCOHAMLPVLTNHIRYHOCMLMDKLGITVRODRCL 893
Db 1 MREVTVELSSQGFMTGIRSDVCOHAMLPVLTNHIRYHOCMLMDKLGITVRODRCL 60
Qy 894 QLAMTHPSHHLNFGMNPDHARNSLSNCGIRPKYGDRAVHNMNRKKGINTLINIMSRIG 953
Db 61 QLAMTHPSHHLNFGMNPDHARNSLSNCGIRPKYGDRAVHNMNRKKGINTLINIMSRIG 120
Qy 954 QDDPPRSRINHEERLEFLGDVAVEFLTSYHLYLPSSLSEGGIATYRTAIVQNOHMLA 1013
Db 121 QDDPPRSRINHEERLEFLGDVAVEFLTSYHLYLPSSLSEGGIATYRTAIVQNOHMLA 180
Qy 1014 KKLLEDPFMYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLEBAKOLFGRLLFNDDP 1073
Db 181 KKLLEDPFMYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLEBAKOLFGRLLFNDDP 240
Qy 1074 LREVMANYPHPIQLOEPNTDROLIETSEVLOKLTPEFEBAIGVIFTHVRLARAFPLRTV 1133
Db 241 LREVMANYPHPIQLOEPNTDROLIETSEVLOKLTPEFEBAIGVIFTHVRLARAFPLRTV 300
Qy 1134 GFNHLTLGNQRMELGDSIMQVATEYFIHFRPDHHEGHTLRLSSLVNNRTQAKVAEE 1193
Db 301 GFNHLTLGNQRMELGDSIMQVATEYFIHFRPDHHEGHTLRLSSLVNNRTQAKVAEE 360
Qy 1194 LGMQSYAITNDKTKRPVGLRTKTLADLLESFIALYTDKDLBYVHTPMANVCFFRPRLKEFI 1253
Db 361 LGMQSYAITNDKTKRPVGLRTKTLADLLESFIALYTDKDLBYVHTPMANVCFFRPRLKEFI 420
```

QY	1254	LNQDMNDKSLQAOCCCLLFRTEGKEPDI.PLYKTT.OTVBSHARTYTVAVFPGEGIGGK	1313
Db	421	LNQDMNDKSLQAOCCCLLFRTEGKEPDI.PLYKTT.OTVBSHARTYTVAVFPGEGIGGK	480
QY	1314	GPSIQQAEMGAMDALEKYNFPQMAHQRFIGRKRYROELKEMRWEREHQEREPEDETDIK	1372
Db	481	GPSIQQAEMGAMDALEKYNFPQMAHQRFIERKRYROELKEMRWEREHQEREPEDETDIK	540
QY	1374	K 1374	
Db	541	K 541	
RESULT 12			
AE21037			
ID	AE21037	standard; protein; 486 AA.	
XX	AE21037;		
DT	11-AUG-2005	(first entry)	
DE	Novel human polypeptide SEQ ID NO 1731.		
XX			
KM	vulnery; CNS-gen.; gene therapy; diagnostic; forensic; mapping;		
KM	DNA purification; protein purification; osteoarthritis; antiarthritic;		
KM	osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;		
KM	periodontal disease; antiinflammatory; mouth disease; burns; injury;		
KM	peripheral neuropathy; Alzheimer disease; neuroprotective; nootropic;		
KM	degeneration; parkinsons disease; antiparkinsonian; neurological disease;		
KM	cerebrovascular ischemia; cerebroprotective; vasotropic;		
KM	cardiovascular disease; autoimmune disease; immunosuppressive;		
KM	immune disorder; viral infection; virucide; infection; cancer;		
KM	cytostatic; neoplasm.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2005049806-A2.		
PD	02-JUN-2005.		
PF	11-MAR-2004; 2004WO-US007412.		
PR	14-MAR-2003; 2003US-00389559.		
XX			
PA	(NUVE-) NUVELO INC.		
PI	Gang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;		
PI	Ghosh M, Xue A, Abund V, Zhao Q, Wang D, Goodrich R, Chen R;		
PI	Wehrman T, Meng G, Boyle B;		
DR	WPI; 2005-417730/42.		
XX			
PT	New polynucleotide encoding a polypeptide with biological activity,		
PT	useful for treating a disease or disorder, e.g. osteoarthritis, burns,		
PT	CNS and peripheral disease, stroke, autoimmune disorders, viral		
PT	infection, or cancer.		
XX			
PS	Example 3; SEQ ID NO 1731; 500bp; English.		
XX			
CC	The invention describes a new isolated polynucleotide (I) encoding a		
CC	polypeptide with biological activity comprising: a nucleotide sequence of		
CC	SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes		
CC	to the sequence of (I) under stringent hybridization conditions; or a		
CC	nucleotide sequence having greater than 9% sequence identity with the		
CC	sequence of (I). Also described are: a(n) (expression)vector comprising		
CC	(I); a host cell genetically engineered to comprise (I) operatively,		
CC	associated with a regulatory sequence that modulates expression of the		
CC	polynucleotide in the host cell; an isolated polypeptide comprising a		
CC	sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide		
CC	18; a polypeptide encoded by (I); or a polypeptide encoded by a		
CC	polynucleotide hybridizing under stringent conditions with any one of SEQ		
CC	ID NOS: 1-567; a composition comprising the polypeptide of (3) and a		

CC carrier, an antibody directed against the polypeptide of (3); a method
CC for detecting (1) in a sample; a method for detecting the polypeptide of
CC (3) in a sample; a method for identifying a compound that binds to the
CC polypeptide of (3); a method for producing the polypeptide of (3); and a
CC collection of polynucleotides, where the collection comprising of at
CC least one of SEQ ID NOS: 1-567. (1) is a polynucleotide comprising any of
CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological
CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:
CC 568-1134. All sequences are fully defined in the specification. The
CC sequences and methods are useful in diagnostics, forensic, and gene
CC mapping, in identifying of mutations responsible for genetic disorders or
CC other traits, in assessing biodiversity, and for producing many other
CC types of data and products dependent on DNA and amino acid sequences. The
CC composition and method are useful for treating a disease or disorder,
CC e.g. osteoporosis, osteoarthritis, periodontal disease, burns, CNS and
CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,
CC autoimmune disorders, viral infection, or cancer. This is the amino acid
CC sequence of a novel polypeptide of the invention.
XX
XX Sequence 486 AA;

	Query March	33.6%	Score 2519;	DB 9;	Length 486;
	Best Local Similarity	99.2%	Pred. No. 4.4e-170;		
	Matches 477;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0
Dy	9894 QIANTHPSHHILNFCGNIPDHAANSLSNCGIRPKYGDGRKYVHHMNRKKGINTLINIMSRIG				953
Db	6 QIANTHPSHHILNFCGNIPDHAANSLSNCGIRPKYGDGRKYVHHMNRKKGINTLINIMSRIG				65
Dy	954 QDDPFSRININERLEFLGDAVVEFLTSVHLVYLPFSLEEGSLATYRTAIVQNOHLAMLA				1013
Db	66 QDDPFSRININERLEFLGDAVVEFLTSVHLVYLPFSLEEGSLATYRTAIVQNOHLAMLA				125
Dy	1014 KKELEDRPMLYAHGPDLCSRESDLRHMANCFEALIGAVYLEGSLBEAKQLFGRLIFNDPD				1073
Db	126 KKELEDRPMLYAHGPDLCSRESDLRHMANCFEALIGAVYLEGSLBEAKQLFGRLIFNDPD				185
Dy	1074 LREVLNLYPLHPLLOEPNTDROLIETSPVLOKLTPEFEALGVFTYHRLARAFLLRTV				1133
Db	186 LREVLNLYPLHPLLOEPNTDROLIETSPVLOKLTPEFEALGVFTYHRLARAFLLRTV				245
Dy	1134 GFNHLTLGHNORMEFLGDSIMQVATETFLFHPDHHGHLLTLRSSLVNNRTOAKVAEE				1193
Db	246 GFNHLTLGHNORMEFLGDSIMQVATETFLFHPDHHGHLLTLRSSLVNNRTOAKVAEE				305
Dy	1194 LGMGEYAITNDKTRPVGLRTKTLADLLESFIALYTDKDLVEYVTEFMNVCFFPRLKEFI				1255
Db	306 LGMGEYAITNDKTRPVGLRTKTLADLLESFIALYTDKDLVEYVTEFMNVCFFPRLKEFI				365
Dy	1254 LNDQMDNPKSLOOCCLTLRTEGKEPDIPLKYTLQTVGSPSHARTYTVAVYFKGERIGGK				1313
Db	366 LNDQMDNPKSLOOCCLTLRTEGKEPDIPLKYTLQTVGSPSHARTYTVAVYFKGERIGGK				425
Dy	1314 GPSIQQAEEMGAMDALEKYNFPQMAHQRFGRKYRQELKEMRWERHQBQREPDETEDIK				1373
Db	426 GPSIQQAEEMGAMDALEKYNFPQMAHQRFGRKYRQELKEMRWERHQBQREPDETEDIK				485
Dy	1374 K 1374				
Db	486 K 486				
RESULT 13					
ID	ADQ96710				
ADQ96710	standard; protein, 466 AA.				
XX	ADQ96710;				
AC					
XX					
DT	23-SEP-2004 (first entry)				
XX					
DE	Human ribonuclease III (RNase III)-like domain.				
XX					
KW	Human; ribonuclease III; RNase III; research purpose; biological purpose;				

KW clinical purpose; cellular interaction.
XX
OS Homo sapiens.
XX
PN US2004126867-A1.
XX
PD 01-JUL-2004.
XX
PF 09-FEB-2004; 2004US-00774974.
XX
PR 06-JUL-2001; 2001US-00900425.
XX
PA (CROO/) CROOKE S T.
XX (WUHH/) WU H.
XX
PI Crooke ST, Wu H;
XX
DR WPI, 2004-516913/49.
XX
PT New isolated nucleic acid molecule encoding human RNase III, useful for
PT research, biological, or clinical purposes, e.g. defining the roles of
PT RNase III and the interaction of human RNase III and cellular RNA.
XX
PS Example 9; SEQ ID NO 37; 31pp; English.
XX
XX The invention relates to human ribonuclease III (RNase III) and its
CC corresponding nucleic acid sequence. The polynucleotide sequence of the
CC invention is useful for research, biological and clinical purposes. It is
CC useful in defining the roles of RNase III and the interaction of human
CC RNase III and cellular RNA. The present sequence is human RNase III-like
CC domain protein.
XX
SQ Sequence 466 AA;

Query Match 32.8%; Score 2463; DB 8; Length 466;
Best Local Similarity 100.0%; Pred. No. 4e-166;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 909 NPDHANSNSGCGIRPKYGDRAVHMHMRKKGINTLINIMSRIGODDTPPSRINNEHL 968
DB 1 NPDHANSNSGCGIRPKYGDRAVHMHMRKKGINTLINIMSRIGODDTPPSRINNEHL 60
QY 969 EFLGDAVVEFLTSVHLVYLPPSLIEGGALATYRTAIVONHMLAKKLEDPFMYAHGP 1028
DB 61 EFLGDAVVEFLTSVHLVYLPPSLIEGGALATYRTAIVONHMLAKKLEDPFMYAHGP 120
QY 1029 DICRESDLHMANCFEALIGAVYLEGSLBEAKOLFGLLPNDPDLREVLNPLPLQL 1088
DB 121 DICRESDLHMANCFEALIGAVYLEGSLBEAKOLFGLLPNDPDLREVLNPLPLQL 180
QY 1089 QEPNTDROLIETSPVLOKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHNORMEF 1148
DB 181 QEPNTDROLIETSPVLOKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHNORMEF 240
QY 1149 LQDSIMQVATEXYLFTHFPDHEGHLLTLRSSLVNRTQAKVAEBLGMQBYAITNDKTKR 1208
DB 241 LQDSIMQVATEXYLFTHFPDHEGHLLTLRSSLVNRTQAKVAEBLGMQBYAITNDKTKR 300
QY 1209 PVGLRKTTLADLIESFTIALYTDKDLLEVYTFMNVCFEPRLKEFTLINQDMNDPKSLOQC 1268
DB 301 PVGLRKTTLADLIESFTIALYTDKDLLEVYTFMNVCFEPRLKEFTLINQDMNDPKSLOQC 360
QY 1269 CULTRTEGKEPDPPLKTLQTVGSHARTYTVAVVFKGERIGCGKPSIQQEMGAAMA 1328
DB 361 CULTRTEGKEPDPPLKTLQTVGSHARTYTVAVVFKGERIGCGKPSIQQEMGAAMA 420
QY 1329 LEKYNPQMAHQKRFGRKYRQELKEMRWERHQEREPEDEDIKK 1374
DB 421 LEKYNPQMAHQKRFGRKYRQELKEMRWERHQEREPEDEDIKK 466

RESULT 14
,AAU20587

ID AAU20587 standard; protein; 378 AA.
XX
AC AAU20587;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human secreted protein, Seq ID No 579.
XX
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiatic; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.
XX
OS Homo sapiens.
XX
PN WO200155326-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001347.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225265P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226811P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0226924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.

PR	08-SEP-2000	2000US-02320601P
PR	08-SEP-2000	2000US-0232061P
PR	12-SEP-2000	2000US-0231968P
PR	14-SEP-2000	2000US-0232337P
PR	14-SEP-2000	2000US-0232338P
PR	14-SEP-2000	2000US-0232339P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0233053P
PR	14-SEP-2000	2000US-0233054P
PR	14-SEP-2000	2000US-0233055P
PR	21-SEP-2000	2000US-0234223P
PR	21-SEP-2000	2000US-0234274P
PR	25-SEP-2000	2000US-0234937P
PR	25-SEP-2000	2000US-0234938P
PR	26-SEP-2000	2000US-0235484P
PR	27-SEP-2000	2000US-0235834P
PR	27-SEP-2000	2000US-0235836P
PR	29-SEP-2000	2000US-0236327P
PR	29-SEP-2000	2000US-0236357P
PR	29-SEP-2000	2000US-0236358P
PR	29-SEP-2000	2000US-0236359P
PR	29-SEP-2000	2000US-0236370P
PR	02-OCT-2000	2000US-0237037P
PR	02-OCT-2000	2000US-0237038P
PR	02-OCT-2000	2000US-0237039P
PR	02-OCT-2000	2000US-0237040P
PR	13-OCT-2000	2000US-0237935P
PR	13-OCT-2000	2000US-0239377P
PR	20-OCT-2000	2000US-0241096P
PR	20-OCT-2000	2000US-0241121P
PR	20-OCT-2000	2000US-0241176SP
PR	20-OCT-2000	2000US-0241766P
PR	20-OCT-2000	2000US-0241787P
PR	20-OCT-2000	2000US-0241808P
PR	20-OCT-2000	2000US-0241809P
PR	20-OCT-2000	2000US-0246177P
PR	08-NOV-2000	2000US-0246171P
PR	08-NOV-2000	2000US-0246474P
PR	08-NOV-2000	2000US-0246475P
PR	08-NOV-2000	2000US-0246476P
PR	08-NOV-2000	2000US-0246477P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246523P
PR	08-NOV-2000	2000US-0246524P
PR	08-NOV-2000	2000US-0246525P
PR	08-NOV-2000	2000US-0246526P
PR	08-NOV-2000	2000US-0246527P
PR	08-NOV-2000	2000US-0246528P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246609P
PR	08-NOV-2000	2000US-0246610P
PR	08-NOV-2000	2000US-0246611P
PR	17-NOV-2000	2000US-0249211P
PR	17-NOV-2000	2000US-0249212P
PR	17-NOV-2000	2000US-0249213P
PR	17-NOV-2000	2000US-0249214P
PR	17-NOV-2000	2000US-0249215P
PR	17-NOV-2000	2000US-0249216P
PR	17-NOV-2000	2000US-0249217P
PR	17-NOV-2000	2000US-0249218P
PR	17-NOV-2000	2000US-0249224P
PR	17-NOV-2000	2000US-0249245P
PR	17-NOV-2000	2000US-0249246P
PR	17-NOV-2000	2000US-0249247P
PR	17-NOV-2000	2000US-0249252P
PR	17-NOV-2000	2000US-0249253P
PR	17-NOV-2000	2000US-0249259P

PR		17-NOV-2000;	2000US-0294300P.
PR		01-DEC-2000;	2000US-0250160P.
PR		01-DEC-2000;	2000US-0250391P.
PR		05-DEC-2000;	2000US-0251030P.
PR		05-DEC-2000;	2000US-0251988P.
PR		05-DEC-2000;	2000US-0256719P.
PR		06-DEC-2000;	2000US-0251479P.
PR		08-DEC-2000;	2000US-0251856P.
PR		08-DEC-2000;	2000US-0251868P.
PR		08-DEC-2000;	2000US-0251989P.
PR		08-DEC-2000;	2000US-0251990P.
PR		11-DEC-2000;	2000US-0254097P.
PR		05-JAN-2001;	2001US-0259678P.
XX			
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA,	Barash SC, Ruben SM;	
DR	WPI:	2001-451931/48.	
N-P	PADB:	AAS33296.	
XX			
PT	New nucleic acids	and polypeptides,	useful for diagnosing, preventing or treating medical conditions.
XX			
PS	Claim 11;	SEQ ID NO 579;	753pp; English.
XX			
CC	The invention relates	to novel isolated nucleic acid molecules (I) encoding human secreted proteins (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression of secreted proteins. (I) and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples, and so which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and thrombosis), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). (I) and (II), agonists, antagonists and antibodies can also be used to promote wound healing, maintain organs before transplantation, and support cell culture of primary tissues.	
Query Match	26.1%;	Score 1961;	DB 4; Length 378;
Best Local Similarity	98.9%;	Pred. No. 1.4e-10;	
Matches 374;	Conservative 0;	Mismatches 4;	Indels 0;
Gaps 0;			
997	ATTTATATGONHLAMAKKLELDPMYLAHPDLCRESDDLHMANACEALIGAVYLEGS	10566	
Db	1 ATTTATATGONHLAMAKKLELDPMYLAHPDLCRESDDLHMANACEALIGAVYLEGS	60	
Qy	LEBAKKOLGRLLPNDPDLREVWLYPLPHLOLQEPNDRQLIEFSPVLOKTEFEBAICV	11164	
Db	61 LEBAKKOLGRLFPNDPDLREVWLYPLPHLOLQEPNDRQLIEFSPVLOKTEFEBAICV	120	
Qy	IFTHVRLARAFRTLVTFGNLTGLGHNOEMFLDSIMQVATRYELFIHPFDHHEGHTLL	11766	
Db	121 IFTHVRLARAFRTLVTFGNLTGLGHNOEMFLDSIMQVATRYELFIHPFDHHEGHTLL	180	
Qy	LRSLSVNRFIOAKVAEBLGMOEVAITNDKTRPVGLRKTTLADLLESFIALYTDKOILEY	12368	

DB 161 LRSLVNNRTOAKVABELGMQSYAITNDKTRPVALRTTTLADLESFIAALYIDKOLEY 240
QY 1237 VHTFNNVCFEPRLAKEPIILNQDWNDDPKSQLOCCCLTLRTGKEPDIPLYKTLQTVGSPSHAR 1296
DB 241 VHTFNNVCFEPRLAKEPIILNQDWNDDPKSQLOCCCLTLRTGKEPDIPLYKTLQTVGSPSHAR 300
QY 1297 TTYTVAVYFKGERIGCGKGPSIQOAEWGAAMDALFKTNFPQMAHQKRFIRKTRQSLKEMR 1356
DB 301 TTYTVAVYFKGERIGCGKGPSIQOAEWGAAMDALFKTNFPQMAHQKRFIRKTRQSLKEMR 360
QY 1357 WEREHOEREPDETDIKK 1374
DB 361 WEREHOEREPDETDIKK 378
RESULT 15
AAU21744
ID AAU21744 standard; protein, 378 AA.
XX
AC AAU21744;
XX
DT 04-DEC-2001 (first entry)
XX
DE Novel human neoplastic disease associated polypeptide #177.
XX
KW Human; neoplastic disease associated polypeptide; cancer;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasoconstrictive.
OS Homo sapiens.
XX
PN WO200155163-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001358.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227099P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229399P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0233964P.
PR 14-SEP-2000; 2000US-0233979P.
PR 14-SEP-2000; 2000US-0233989P.
PR 14-SEP-2000; 2000US-0233999P.
PR 14-SEP-2000; 2000US-0234009P.
PR 14-SEP-2000; 2000US-0234011P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237040P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246539P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.

PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249246P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.

PI Rosen CA, Barash SC, Ruben SM,
XX WPI, 2001-465558/50.
DR N-PSDB; AAS34943.
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid arthritis
XX
XX Claim 11; SEQ ID NO 471; 687p; English.
XS

The present invention relates to the isolation of novel human neoplastic disease associated polypeptides, and cDNA (AA354767-AA355050) and DNA sequences encoding for these polypeptides. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of disorders involving neoplastic disease such as hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem glioma, adult liver cancer, childhood cerebellar astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may also be useful for treating other disorders such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders and renal disorders. The polynucleotide sequences of the invention are also useful in gene therapy. AAU21568-AAU21851 represent the novel human neoplastic disease associated polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published.pct_sequences

SQ Sequence 378 AA;

Query Match	26.1%	Score 1961	DB 4	Length 378
Best Local Similarly	98.9%	Pred. No. 1.4e-130		
Matches 374	Conservative 4	Mismatches 0	Indels 0	Gaps 0

QY 997 ATYPTAIQONHMLAKKLEIDPMLYAHGDI.CRESDLRHMANCEALIGAYLEGS 1056
 |||||
 Db * 1 ATYPTAIQONHMLAKKLEIDPMLYAHGDI.CRESDLRHMANCEALIGAYLEGS 60
 |||||
 QY 1057 LEEAKQLFGRLLFNDPDLREVWLNPLPLQIQEENTROLLENSPVQLKTEFEAGV 1116
 |||||

Db	61	LEAKOJFGLLFFNDPOLREVLINVPALPIQOEPNTDROLIETSPIVLOKLTREEEBAIGV	120
Qy	1117	IFTHVRLAAAFILTRITGFMHLLTGNHORMEPLGDSIMOLVATEYILTHRPDHHGHLTL	1176
Db	121	IFTHVRLAAAFILTRITGFMHLLTGNHORMEPLGDSIMOLVATEYILTHRPDHHGHLTL	180
Qy	1177	LRSSLVNNRTQAKVAEBELGMOEYAITNDKTRPVGLRTKTLADLSEFIAALYTDKOLEY	1236
Db	181	LRSSLVNNRTQAKVAEBELGMOEYAITNDKTRPVALTRTKTLADLSEFIAALYTDKOLEY	240
Qy	1237	VHTFMNVCFFPRKLKEFILLNDMDNDPKSOLOOCCILTPBEKEPRLPYKTLQWGPSHAR	1296
Db	241	VHTFMNVCFFPRKLKEFILLNDMDNDPKSOLOOCCILTPBEKEPRLPYKTLQWGPSHAR	300
Qy	1297	TYTAAVVFKEBRIIGCGKPSIQQAEMGAAMDALAEKYNPOMAHOKREIIGRTYROELKEMR	1356
Db	301	TYTAAVVFKEBRIIGCGKPSIQQAEMGAAMDALAEKYNPOMAHOKREIIGRTYROELKEMR	360
Qy	1357	WEREHOREPDEPTEDDICK	1374
Db	361	WEREHOREPDEPTEDDICK	378

Search completed: February 3, 2006, 19:36:16
Job time : 219 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using BW model

Run on: February 3, 2006, 19:36:34 ; Search time 55 Seconds
(without alignments)
2403.670 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 NM0GNTCHNMFHGRGCRP.....MRWERHOREPDETDIK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	890	11.9	412	2	T21419	hypothetical prote
2	558	7.4	682	2	T21420	hypothetical prote
3	316	4.2	551	2	S57447	HPBRI-7 protein -
4	315	4.2	249	2	B69693	ribonuclease III (
5	307.5	4.1	229	2	AE1300	ribonuclease III (
6	300	4.0	229	2	AE1672	ribonuclease III h
7	293	3.9	1560	2	T42727	proliferation pote
8	285.5	3.8	225	2	B82073	ribonuclease III v
9	282	3.8	263	2	A83961	ribonuclease III r
10	281.5	3.8	232	2	P98012	ribonuclease III (
11	279.5	3.7	232	2	H95144	ribonuclease III (
12	276	3.7	948	2	A57640	retinoblastoma bin
13	275	3.7	760	2	T06291	extensin homolog T
14	274.5	3.7	1531	2	T48946	hypothetical prote
15	273.5	3.6	1006	2	G86292	hypothetical prote
16	273	3.6	891	2	G84693	probable prolina-r
17	271.5	3.6	620	2	S06733	hydroxyproline-ric
18	270	3.6	226	2	H84959	ribonuclease III (
19	268.5	3.6	231	2	B84725	ribonuclease III (
20	267	3.6	224	2	A81260	ribonuclease III (
21	267	3.6	243	2	D89896	RNase III (importe
22	267	3.6	998	2	T30930	hypothetical prote
23	265.5	3.5	1002	2	T43336	carboxypeptidase C
24	262	3.5	226	1	NRE63	ribonuclease III (
25	262	3.5	226	2	P85902	RNase III, de RNA
26	262	3.5	226	2	A91058	extensin precursor
27	259	3.5	322	2	S25299	hypothetical prote
28	258	3.4	231	2	S76204	hypothetical prote
29	257.5	3.4	272	2	T35656	ribonuclease III -

30	256	3.4	246	2	H70187	ribonuclease III (
31	255.5	3.4	1638	2	A42091	transcription acti
32	255	3.4	226	2	AC0829	ribonuclease III (
33	254	3.4	1611	2	T38236	hypothetical prote
34	253.5	3.4	368	2	C29356	hydroxyproline-ric
35	253.5	3.4	1048	2	T31425	C-terminal domain-
36	252	3.4	226	2	AF0331	ribonuclease III (
37	251.5	3.4	240	2	F64602	ribonuclease III (
38	251.5	3.4	1201	2	G84534	unknown protein (1
39	250.5	3.3	727	2	C84534	hypothetical prote
40	249	3.3	225	2	T09964	extensin CYC15 pre
41	249	3.3	907	2	E96836	hypothetical prote
42	249	3.3	1877	2	T21861	hypothetical prote
43	245	3.3	230	2	G97115	dsRNA-specific rib
44	244.5	3.3	489	2	T11622	extensin class 1 p
45	243.5	3.2	708	2	D96711	hypothetical prote

ALIGNMENTS

RESULT 1

T21419
hypothetical protein F26B4.13 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_rev15on 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T21419

R/Highlighting, J.
submitted to the EMBL Data Library, October 1996

A/Reference number: Z19419

A/Accession: T21419

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-412 <NLT>

A/Cross-references: UNIPARC:UPI000017B936; EMBL:Z81070; PIRN:CAB03005.1; GSPDB:GN00019; (

A/Experimental source: clone F26B4

C/Genetics:

A/Gene: CESP:F26B4.13

A/Map position: 1

A/Introns: 359/2; 400/3

Query Match	11.9%	Score 890;	DB 2;	Length 412;
Best Local Similarity	43.1%	Pred. No. 6.6e-40;		
Matches 181;	Conservative 85;	Mismatches 136;	Indels 18;	Gaps 7;
QY	944	TLINIMSRIGQDDPTSRINHNRELEPLGDVAVEFLTSVHLYLTFPSLBEGLATRTAI	1003	
DB	2	SLFNIMKGTSGGP---ILHNERLEYLDVAVELVSHLYFMLTHFEGGLATYRTAL	57	
QY	1004	VONQHLAMLAKELEDPFMYAHGPDLGRESDLRHMANCFEALIGAVYLEGSLBEAKOL	1063	
DB	58	VONRNLAATLAKNCRIDEMLOYSHGADLINVABEKALIANFEVMAIYIDGGLACDVI	117	
QY	1064	FGRLIF-NDPDLREVLNYPPLHPLQLOBPNTDROLIETSPVLOKTEFEAIGVIFTHVR	1122	
DB	118	FSKAWYGHQPVLEKEKDHINEHELEKREDPGDRDLSFIPPTLSTFPALEBRILQFNMR	177	
QY	1123	LARATLTATVGNHNLTLGHNOEMFLGDSIMQVATVETLFIHPPHHGHCHTLTLASSLV	1182	
DB	178	LAKAFTRRNIPNDTLKGNORLEWLGDSVLOLVSDFLYRFPFHGHGMSLTLSTLV	237	
QY	1183	NNRTOKVAEELGMOEYATINDTKRPV---GLRTKTLDDLSEFTAAALYTDDLEAVHT	1239	
DB	238	SNQTOAVVCDGGLGFTFV-----KAPYTPPELKDKADLVAFGALYVDRGIEHCRA	292	
QY	1240	FNWVCFPPRLKEPILNQDNNDPKSOLQOCCLTLR-TEGKEPDIPLYKTLQTVGPHARTY	1298	
DB	293	FRIVFCPLRKPISEKKNDAKSHLQWCCLMARDDSSSEDPMEYRVIGISGPTNRIIF	352	
QY	1299	TVAVYFKGRIRIGCGKPSIQOAEWGA---AMDALEKYNPPOM-AHQKRTIGKRYRELKE	1354	
DB	353	KIAVYVYKGRILASPAESNVYKAKELRYAELLALNLTESMSPSKKAKKNSNMRRLEQDTS	412	

```

RESULT 2
T21420
hypothetical protein F26E4.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21420
R:Lightning, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19419
A:Accession: T21420
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-682 <MIL>
A:Cross-references: UNIPARC:UPI0000179272; EMBL:281070; PIDN:CAB03006.1; GSPDB:GN00019;
A:Experimental source: clone F26E4
C:Genetic:
A:Gene: CESP.F26E4.10
A:map position: 1
A:introns: 119/3; 226/2; 303/2; 377/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F26E4.10

Query Match
Best Local Similarity 7.4%; Score 558; DB 2; Length 682;
Matches 178; Conservative 126; Mismatches 296; Indels 140; Gaps 23;

Qy 277 RHRSYRSRERERHRHRDNRSPSLERSYKRSGRVGLSVVPAGCTPELP-G 335
Db 14 KHKRARKKKYQKQYGRHKEEMWQQLGRFRQNPSTSSAPDTEKILPTESTSLPFG 73
Qy 336 EIIKNTDS-----WAPLEIYNHRSPSREKGRAMWEERKMSDNOSSGDKNTSTK 388
Db 74 DSERLTEKQYETNYMIDP--VVSTHSALIK-----SNRVLIK 110
Qy 389 EKPEETMPDKNEESEEELK-----PWIRCTHS----- 418
Db 111 ABEAEKYMMIKAKSTSKILQDQTKILEVTKRKLQADVPYIIPCHSMKGRKTPKOK 170
Qy 419 ---ENYSSDPMDQVDSITVGTSLRLDLVKFESEELSGROEKAKARPP--WEPPKYL 473
Db 171 GGESEFTASVDSDSNDS-----QDEASISEPTNRQAPRADKT 208
Qy 474 DEDELESSSESCSDSDSTCS-----SSDSEVNDVLAETIKKKAHADRHL 519
Db 209 GE-----VKDEKQTCNRRNQORAKRLNFEKEKQITLLKKGIDRKKTHPNGIH 258
Qy 520 DELWYNDPGOMNDGPCKCSAKARTGIRHSIYGEBAIKPCRPMTNNGRLPHY--RIT 577
Db 259 PDISFNEKGLGNBGEPCRCPEPIKTCGLKRGYAGEDKALDCK--KSNGENLHYTLRYT 316
Qy 578 VSPPTNFLTDRPTVIEYDHEYLFEGFSPFAHAPL---TNIPLCKVIRPNIDYTHITE 633
Db 317 PLPSENGQL--YRTHMAINGEPEFEGRSLTHAPLPDCTRABICK--YSMDYEPOLVE 371
Qy 634 EMMP--ENPCYKGLSLPLFRDILLYDMNLKGLPEDSPPCCRPFHMPRVRLPFG 692
Db 372 EFPNPDCPPEEDDMLEFYIFHEIFEMDLFELPKHI PSDVESCPMIIHMPRVQ--TKID 430
Qy 693 GKEVLVMOHQLYLL--RCSKALVPEEELANMLQWELEMOKYAECKGMIVTPGKPPSS 751
Db 431 LVQUMSSKTVLAYFTSKSGSEIISPEDVNRICSAQIDQFRNTSKKQSI VLTNKKRPSA 490
Qy 752 VRIDQLDRQFNPDVITPPIIVHFGIRPAQLSYAGDPQYOKLMSYVKLPHLLANSPKYK 811
Db 491 IRADWERDEBEKKEV---YVHNALRAQYTAISLPRIAFLEKTLNKMIQ--EKQSSGYV 545
Qy 812 QTDKQLAEEREALQKIRQKNTKREYVTELSQGWKIGIRSDVOCHAMMLPVLTNHR 871
Db 546 NKDFEK---TKMELFHLKRENSARMLKIREPVAIGEITGLKEDVAHVMTTLACHIR 602
Qy 872 YHQCGLMLDKLIGYTPQDRCLDLAMTHPSHHNLFNGNPNPHANSISNGCIRPKYGDK 931
Db 603 YNFSLVPEFEVLEIKFNDRKRVIELALMHSSFKSHYGTPIDVHKMINTNGYRR-KTG--- 658

```

```

Qy 932 VHHNMRKKKGINTLINMSR 951
Db 659 AEDKREKKRGRSLKUKFLDK 678

RESULT 3
S57447
HPRIT-7 protein - human
N:Alternate names: HPRIT-4 protein
C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Oct-2004
C:Accession: S57447; S57489
R:Feischauer, K.L.
submitted to the EMBL Data Library, June 1992
A:Reference number: S57447
A:Accession: S57447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <FLB>
A:Cross-references: UNIPROT:Q16630; UNIPARC:UPI000006D566; EMBL:X67336; NID:9871300; PIDN:
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-551 <FL2>
A:Cross-references: UNIPARC:UPI000006D566; EMBL:X67337; NID:9871298; PIDN:CAA47752.1; PII
A:Genetic:
A:introns: 231/3
F:82-151/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match
Best Local Similarity 4.2%; Score 316; DB 2; Length 551;
Matches 111; Conservative 32; Mismatches 106; Indels 148; Gaps 21;

Qy 19 PRGCGHGAAPSPAFRRPQNLRLHPQRPVQYQYEPSPASTTSSDPAPNF---LPP 74
Db 198 PGCGRGRRGRRPFA-----VPGGDRRPGRRAGP-----CGRRPPPPAGQTRP 237
Qy 75 RPDVPPPPMPMPASAGLPRLPCPIRPP---FPNHOMRHPPV-----PPC--FP 119
Db 238 RP---PLGPPGPPGPPGPPPPGQVLPPLACGPPNKGDRPPVPLPFGQPPGQPLGLRPP 294
Qy 120 MPPWPCPNPPVPAPAPQO-----GFPF-----FMPPPPMPPPP--PPV 159
Db 295 GPPP--PVEGYPGPPGPPPPGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 353
Qy 160 MFOQVYQYPPGYSHHNPPEPFNSFQNNPSSFLPSANNSSP--HFRHLPPY----- 210
Db 354 AF-HYN---PAF-----PPPP-----NSGMPITSDSRGPPPPDPYGRPPPYDRGDYV 396
Qy 211 PLPKAPSERRSPERLKHVDHHRHDS----- 237
Db 397 PGRREMDTARTPLSEAEFEIEMNRRAISSSAISRAVSDASAGDYSAIETLVTAISLK 456
Qy 238 -----HG-RGHRHSIDREBRGSPRRARODSRRSDYDRGRTP 275
Db 457 GSKVADDRCKYLISLDDCLHGISTSKYSGSSRRBSRERHDSR--SREKSRHKSRR 514
Qy 276 SHRSYRSRERERHRHRDNRSPSLERSYKRYK 312
Db 515 DRHDDYRRSRERERHRHRDNRD--RDRDRDRDRERER 549

RESULT 4
B69693
ribonuclease III (EC 3.1.26.3) - Bacillus subtilis
N:Alternate names: RNase D; RNase O
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
R:Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero,
C.; Bron, S.; Brouillet, S.; Bussch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi,
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Evington, J.; Fabret, C.; Ferrari, E.

```

Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fume, S.; Galizzi, A.; Gallen
Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeber, P.; Koningsstein, G.; Krog, S.; Kumano, M.; Kurita, K.; Lapilus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maseel
Y.M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schlach, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B63693
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-249 <KUN>
A:Cross-references: UNIPROT: P51833; UNIPARC:UPI0000060300; GB:Z99112; GB:AL009126; NID:G
A:Experimental source: strain 168
R:Oguro, A.; Kakeshita, H.; Takamatsu, H.; Nakamura, K.; Yamane, K.
Gene 172, 17-24, 1996
A:Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit, on
A:Reference number: J04819; MUID:96257247; PMID:8654983
A:Accession: J04821
A:Molecule type: DNA
A:Residues: 1-78, 'P', 80-137, 'K', 139-170, 'P', 172-242, 'E', 244-249, 'INPPYDSGGFQVVCRLI' <OGU
A:Cross-references: UNIPARC:UPI00001759DE; DDBJ: D64116; NID:G1389548
A:Note: the sequence in Genbank entry D64116, release 106.0, (PID:G1389549) has 138-Glu
C:Genetics:
A:Gene: rncs; srb
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
C:Keywords: hydrolase
F.172-245/Domain: double-stranded RNA-binding repeat homology <DSR>
Query Match 4.2%; Score 315; DB 2; Length 249;
Best Local Similarity 34.7%; Pred. No. 8.7e-10;
Matches 85; Conservative 46; Mismatches 96; Indels 18; Gaps 7;
QY 1104 LOKLTFEFAIGVIFTHVRLAARAFTRVGFNHLTG--NORMFELDSIMQVATEY 1161
DB 17 VQPFQFQRIHVHFNQEKLYQAFTHSSYVNEHKKYFEDNEHLEFLDPAVLELTISRF 76
QY 1162 LFIHPDHNGHLLTLRSSLVNNRTOAKVAELGMOEYAITNDKTRPVG--LRTKTLAD 1220
DB 77 LFAKPRMSEGLTLTKRAIVCEPSLVSLAHELSEFDVLVLGKGEEMTGRRPALADV 136
QY 1221 LESFIALYTDKLEVHTFMVNCFFPRKLEFLLNDMDNDPSQLQOCCLTARTGSK--- 1277
DB 137 FFAFGALYLDQGLFVSEFLLKVVYFPRKINDGAFSHVM--DFXSQLOE---YVORDKGS 192
QY 1278 EEDIPLYKLTQVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEKYNFPM 1337
DB 193 E-----YKTSNEKGPANREFAIVSLKGPGLGVGNGKKEAQAADALAKL---EK 244
QY 1338 AHQKR 1342
DB 245 HHTKQ 249

A:Accession: AE1300
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <GLA>
A:Cross-references: UNIPROT: O8Y691; UNIPARC:UPI00000552CB; GB:NC_003210; PIDN:CAC99883.1;
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: rncs
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
Query Match 4.1%; Score 307.5; DB 2; Length 229;
Best Local Similarity 33.0%; Pred. No. 2e-09;
Matches 76; Conservative 51; Mismatches 86; Indels 17; Gaps 7;
QY 1109 EFEAIGVIFTHVRLAARAFTRVGFNHL--TLGHNQMEFLGDSIMQVATEYLFHF 1166
DB 6 ELQESVGFDFQVVELLKQAFTHSSYVNEHRRNVNDNELEFLGDAVLELTVDYLFNFKY 65
QY 1167 PDHHEGHLTLRSSLVNNRTOAKVAELGMOEYAITNDKTRPVGLRTK--TLADLLESF 1224
DB 66 PMAEGHMTKMAAIVCEPSLVEFAVAFSKY-VRLGGEERKAGGRTPALADVFESEF 124
QY 1225 IAAITVDKDLERYVHTFMVNCFFPRK--EFLNODMNDPSQLQOCCLTARTGKEPDIP 1282
DB 125 ICAIYLDNGIDKRVTFLEKRVTPKIDAGAYLQTVDY--KTLDQETV-----QRDRDV 175
QY 1283 L-YKTLQTVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEKYNFPM 1331
DB 176 IEYDILGERTGPANRKAFAQVAVVNGVVLGKSGRTKQAEQAAPALAK 225

RESULT 6
AE1672
ribonuclease III homolog rncs [imported] - *Listeria innocua* (strain Clijp11262)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1672
R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
A:Authors: Krefl, J.; Kuhn, M.; Kunat, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genome of *Listeria species*.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <GLA>
A:Cross-references: UNIPROT: O92AK3; UNIPARC:UPI0000134353; GB:AL592022; PIDN:CAC97149.1;
A:Experimental source: strain Clijp11262
C:Genetics:
A:Gene: rncs
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
Query Match 4.0%; Score 300; DB 2; Length 229;
Best Local Similarity 31.8%; Pred. No. 4.9e-09;
Matches 76; Conservative 53; Mismatches 88; Indels 22; Gaps 8;
QY 1109 EFEAIGVIFTHVRLAARAFTRVGFNHL--TLGHNQMEFLGDSIMQVATEYLFHF 1166
DB 6 ELQESVGFDFQVVELLKQAFTHSSYVNEHRRNVNDNELEFLGDAVLELTVDYLFNFKY 65
QY 1167 PDHHEGHLTLRSSLVNNRTOAKVAELGMOEYAITNDKTRPVGLRTK--TLADLLESF 1224
DB 66 PMAEGHMTKMAAIVCEPSLVEFAVAFSKY-VRLGGEERKAGGRTPALADVFESEF 124
QY 1225 IAAITVDKDLERYVHTFMVNCFFPRK--EFLNODMNDPSQLQOCCLTARTGKEPDIP 1282
DB 125 ICAIYLDNGIDKRVTFLEKRVTPKIDAGAYLQTVDY--KTLDQETV-----QRDRDV 175
QY 1283 L-YKTLQTVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEKYNFPM 1340
DB 176 IEYDILGERTGPANRKAFAQVAVVNGVVLGKSGRTKQAEQAAPALAK 225

Db 176 IEYDLGETGPAHKAFAQAQIVNGOVLGSGSRTKKAQEGSAQAFAIN-----QUTHR 229

RESULT 7

742727

Proliferation potential-related protein - mouse
C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42727

R:Mitte, M.M.; Scott, R.E.
submitted to the EMBL Data Library, November 1998

A:Reference number: Z22246

A:Accession: T42727

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1560 <M17>

A:Cross-references: UNIPROT:P97868; UNIPARC:UPI0000028767; EMBL:U83913; NID:G9358884; PI

A:Experimental source: strain Bald/C

C:Gene: p2p-R

C:Function: A:Description: Involved in hnRNP association and Rb1 binding

F:57-107/Domain: RING finger homology <RRN>

Query Match 3.9%; Score 293; DB 2; Length 1560;
Best Local Similarity 23.8%; Pred. No. 1.3e-07;
Matches 151; Conservative 72; Mismatches 200; Indels 212; Gaps 34;

25 HBARSA--PSFRPQLRLHQQPPVOGYEPSPA-----PSTFSSPAVNPFP 73

296 NARPEGGPGWBNH-N-KLGYLVSPPOQLRGERSTYRSINGRHHSERSQRTQDS-IP 353

74 PRPDVFPFPPPPPSAQGLPCCPIRPPPNQMRHFPVPCFPMPMPMCNNPVP 133

354 ARPVFPVPVPP-----ELVPPPP-----HTLPLP--PCVLPFGSPQPP--P 391

134 GAPPGCGTFPPMPPPSMFPP-----PPVMPQCVNT----- 166

392 GQPPTAG--YAVPPPGFPAPANISTACGPGVPTAHSNMTPTTQAPLISBEFYREOK 448

167 -----QVP--PGVHNHFPFPPSPFSPQNNBSPFSAANNSSPHFRLLPPPLPKAS 217

449 RUKESKFPYSSGSSYSRSSTYDSKSRSGSTRSRYSRSFSKSHSISYSPPT----- 502

218 ERSRPERLGHYDDHRRDHSHGGERHRSLLDRERGRSPDRRDSRYSDYD-RGRTPS 276

503 -RRGRCKSNY--KRSRSHG--YHRS-----KRSPPYRRYHRSRSPQAFRGQST 550

277 RHRSYERSRERER-RRHR-----DNRSPSLERSYKKEYR 313

551 K-RNVPRG-ETEREYFNRYREVPPPYDIKAYGVSVDFRDPFEKERYREMERKYREMYEK 608

314 SGRSYGLSVVPEPACTPPLPGIINKTSMAP-----PLEIVNHRSPSEKRRAREEEK 369

609 YKGVAVGAQPRSA-----NRDPSPERLLPLNIRN--SPFTRGRREDYAAQ 655

370 DRWSDN-----QSSGDKNTYSIKEKEPEETMPD-----KNE---EE 403

656 SHRNRLVGNYPKULSTRSHNAKNPKSKKESEVAVPDGKGNKHKHKKRKNKEKGE 715

404 EEBLLKPVI---RCHSENY-----YSSDMDOVGDSVTVGTSR 440

716 SSSFLNPBELLGPKRCRGSSGIDETKTDTLFFVPPSHEDATPVVDEBMD--ABSITPKSVS 773

441 LNDLVKFEELGSRBEKAKA-----APPPEPKTKLDELLESSEB---CES 487

774 DDXDKRKDPKPKVSDTKRKSDGSAATAKONVLPKSGQEKVGDGKREXPSREPLPKKA 833

488 DEDST-----CSSSSDEVPDVIATIKRKAH 514

834 KEATKIDSVKPSSSSQKDEKVTGT-----PRKAH 863

RESULT 8

B82073

Ribonuclease III VC2461 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82073

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardoon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: B82073

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-225 <HE1>

A:Cross-references: UNIPROT:Q9KPB2; UNIPARC:UPI0000134369; GB:AE004316; GB:AE003852; NID:

A:Experimental source: serogroup O1, strain N16961, biotype El Tor

C:Gene: VC2461

C:Superfamily: 1
A:Map position: 1

Query Match 3.8%; Score 285.5; DB 2; Length 225;
Best Local Similarity 32.2%; Pred. No. 2.8e-08;
Matches 75; Conservative 42; Mismatches 101; Indels 15; Gaps 5;

1101 SPVLQKTEFEBAIGVIFTHVLLARAFLLRTVGFNNHLLTGHNMREFLSDSIMOLVATE 1160

2 TPNMKNLT---SKGYTRETETLNLALTTHRSANK-----HNERLEFLGDSILSFVID 53

1161 YLFHPDHHGHLLTLRSSLVNNRTQAKVAEELGMOEATITNDTKRPVGLRTKT-LAD 1219

54 ELYRRFPKVNEDMSRMATLVRGNTLAELEGFPLGDTLKLGPBELKSGGFRRSIAD 113

1220 LLESFIALYTDKOLEVYHTEFNVCFPPLPKFILLNQMDNDPKSLOQCCLLRTGEXEP 1279

114 AVEALITGALYDSDETARSIVLEWYHGLEIKGASQDKTKRIQEF-----LQGRKK 168

1280 DIPLYKTLQTVGPHARTYVAVVEFG-ERICGCGPSIOQAEWGAMDALEK 1331

169 PLPVYTVNINKEANQERTVACEVAGMDTPVIGKTSRKRAEQAAALRALQ 221

RESULT 9

A83961

Ribonuclease III rncs [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: A83961

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83960; MUID:20512582; PMID:11058132

A:Accession: A83961

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <STO>

A:Cross-references: UNIPROT:Q9KA05; UNIPARC:UPI0000134340; GB:AP001515; GB:BA000004; NID:

A:Experimental source: strain C-125

C:Gene: rncs

C:Superfamily: 1
A:Map position: 1

Query Match 3.8%; Score 282; DB 2; Length 263;
Best Local Similarity 30.9%; Pred. No. 5.2e-08;
Matches 77; Conservative 50; Mismatches 110; Indels 12; Gaps 5;

1084 HPLQLEPNTDQLTETSFVLQKTEFEBAIGVIFTHVLLARAFLLRTVGFNNH--LTIG 1141

15 HSERRRQP---KRLTLTKAQGMFDELRLTLNLTFANKKLVQAFTHSSYVNEHRIOSCK 71

1142 HNQRMEFLGDSIMOLVATVEYLFHFPDHHGHLLTLRSSLVNNRTQAKVAEELGMOEYAI 1201

Db 72 DNERLEFEGDVLAVELANSQVLYKAFBQMSBEDMDKTLRASTICEPSLAQIAEBELFGEVLV 131

Qy 1202 TNDPTKGPVVG-LRTKTLADLSEFIALYTDKDELYVTFMNVCFPRLKEFILNDPMD 1260

Db 132 LGKSEWMTGGRKRALADVAFESVGNALYLDQGNDAVYLFLERITYPISGASH-NMD 190

Qy 1261 PKSOLQOCCLLTRTEGKEPDIPLYKTLQTVGSPSHARTYTAAYFKGERIGCGKPSIQQA 1320

Db 191 FKSOLOEF-----IQDNLIGHIHEYIVOERGPANRREVSIVLNNETLIGVGTGRSKXA 245

Qy 1321 EMGAAMDAL 1329

Db 246 EQHAAQAL 254

RESULT 10

F98012

ribonuclease III (EC 3.1.26.3) [imported] - Streptococcus pneumoniae (strain R6)

C/Species: Streptococcus pneumoniae

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C/Accession: F98012

R.Hookins, J.A.; Albhorn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

E., R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y., P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellidj, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:2143245; PMID:11544234

A:Accession: F98012

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-232 <KUR>

A:Cross-references: UNIPROT:Q8DPJ8; UNIPARC:UPI0000E35D9; GB:AE007317; PIDD:AAK99930.1,

C:Genetics:

C:Gene: rncs

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 281.5; DB 2; Length 232;

Best Local Similarity 32.9%; Pred. No. 4.7e-08;

Matches 79; Conservative 47; Mismatches 91; Indels 23; Gaps 9;

Qy 1099 ETSVPLQKLFEEBAIGVIFTHVRLARAFLLRTGVGNH--LTGHNQRMEEFLDSTIMOL 1156

Db 3 ELQVTLKNHFEI-----FADKKLLETAFTHTSYANEHRLKISHNERLEFLDAVILQ 56

Qy 1157 VATEYLFTHFPHDHNHGHLLTRSSLVNNRTPQAKVAEELGMEVATINDKTRPGLATK- 1215

Db 57 LISSEYLVKKYKKKEGGDSLKRLAMIVREESLAGFARDCQDFQ-IKLGKGEKSGGNRD 115

Qy 1216 -TLADLSEFIALYTDKDELYVTFMNVCFPRLK--EFLINDPMDPKSOLQOCCLLT 1272

Db 116 TILDAAEAFGALLLDKDAKVEFIQVNIIPKVEAGEFEMITDY--KTHLDE--LL 169

Qy 1273 RTEGKEPDIPLYKTLQTVGSPSHARTYTAAYFKGERIGCGKPSIQQAEMGAAMDAL 1331

Db 170 QVNG---DVAIRYQVISTGPADHKVFDVEVLVGEKSGIGQSGRSKKLAEQAANKVAE 226

RESULT 11

H95144

ribonuclease III [imported] - Streptococcus pneumoniae (strain TIGR4)

C/Species: Streptococcus pneumoniae

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C/Accession: H95144

R.Hookins, J.A.; Albhorn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,

neon, T.; Hickey, E.K.; Holt, I.E.

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A>Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: H95144

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <KUR>
A:Cross-references: UNIPROT:Q970G6; UNIPARC:UPI0000051783; GB:AE005672; PIDN:AAK75353.1;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1248
C:Superfamily: ribonuclease III, double-stranded RNA-binding repeat homology

Query Match 3.7%; Score 279.5; DB 2; Length 232;
Best Local Similarity 32.5%; Pred. No. 6e-08;
Matches 75; Conservative 48; Mismatches 91; Indels 17; Gaps 8;

OY 1108 TEFEBAIGVIFTHVALLARAFILRTVGNN--LTGHNQMEFLGDSIMOLVATEYLFIH 1165
 :
DB 6 TVLKNHPAIEFADKKLETAFTHTSYANEHRLIKISHNERLEFGAVLDLLISELYLK 65

OY 1166 PPDHNEGLTLRSSLVNNRTOAKVAEELGMGEVALTNDTKRPVGRTRK--TLADLES 1223
 :
DB 66 YPKKEGGDLSIKRAMIVREESLAGFARDCCPDFP--IKLGKEKSGGRNRDTIIGDAFEA 124

OY 1224 FIALVTYTDLEVTYFMNVCFEPRLK--EFILNQDMNPKSQLOQCCLTLRTGKEKEDI 1281
 :
DB 125 FLGALLDKDVAKVEFEIQWMI PKVEAGEFENITDY---KTHLQE---LLQVNG---DV 175

OY 1282 PL-VYTLQTGVPSHARTYTVAAYFKGERIGCGKGPISIQQAEMGAMDALEK 1331
 :
DB 176 AIRYGVISETGPADHKVFDEVLVEGSKIGQGGRSKLAEOEAANKAVEK 226

RESULT 12

A57640
retinoblastoma binding protein RBQ-1 - human
C:Species: Homo sapiens (man)
C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C:Accession: A57640
R:Sakai, Y.; Saito, M.; Coelho, K.; Kishino, T.; Nakawa, N.; Taya, Y.
Genomics 30, 98-101, 1995
A>Title: cDNA sequence and chromosomal localization of a novel human protein, RBQ-1 (RBBI
A:Reference number: A57640; WUID:96129310; PMID:8595913
A:Accession: A57640
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-948 <SAK>
A:Cross-references: UNIPROT:Q15290; UNIPARC:UPI000006DBAD; GB:X85133; MID:g728590; PIDN:C
C:Genetics:
A:Gene: GDB:RBBP6
A:Cross-references: GDB:626076
A:Map position: 16p12-16p11.2
C:Keywords: tandem repeat; zinc
F;79-129/Domain: RING finger homology <RNG>

Query Match 3.7%; Score 276; DB 2; Length 948;
Best Local Similarity 23.0%; Pred. No. 5.5e-07;
Matches 149; Conservative 61; Mismatches 208; Indels 230; Gaps 31;

OY 27 ARPSA--PSFRQNRLHPQQRPVQGYXEPSSA-----PSTTFSNSPAPNFPR 75
 ||| | :
DB 320 ARPGRGRGHEMSN-KLGLVLPSPQDIRGERSCHYSINRGHNHSRSORTGFS-LPAT 377

OY 76 PDFVPFPMPMPSSAQRLPRCPRIRRPFNNQMHRPPVRPCFFMRPMRCNNPVVGA 135
 :
DB 378 PVFVLPVPPP-----PLVPPP--HTLPLP---PGVPPPOFPSQGF--PGQ 415

OY 136 PRGGCTFRMMMPRSMRPPRRPVMPQVN-----YQYR-- 170
 ||| | :
DB 416 PPPAG--YSVPPGPP-PAPANLSTPWVSGVQTASHNTIPTTQAPRLSRBEFYRQR 471

OY 171 -----GYSHNNEPPSFSNS 184
 ||| | :
DB 472 LKEEEKKSKLDEFNDPAKELMEYKKIQERRRSRSRKSPYSSGSYSASSYSTYKSR 531

OY 185 FÖNNSSFLPSANNSSPHFRLLPYPLPKADEBRRSPERLIKHYDDHRRDHSHGGERH 244

Db 532 GSTRRSYVSRSFSRSRSYSRSPYP-----RRGRGKSHNY--RSRSRSHG---YH 578
QY 245 RLSDLRERRESPDRRQDSRYSDVD-RGRTPSRHYSRSHRER----- 289
Db 579 RS-----RSRSPRYRHSRSRSPQAFRQSPVKRNVPGETREYRNYREVPYDMK 633
QY 250 -----ERRHRDRNRSPELERSYKYEKRSGRSYGLSVPEPACGTBELPGE 336
Db 634 AAYGASVDRPDPEKERYNE-----WEKKYREWEKYYKGYAAGAQPSPA----- 679
QY 337 ITKATDSAP-----PLEIYNHRSPESEKQKARWEEKDMWSDNOSG-----KD 381
Db 680 ---NNEFNPERFLPLINIRN--SPFTRGRREDYVGQSHRSNRISGNYPEKLSARDGHNQ 734
QY 382 KNYTSIKEKEPETWMD-----KNEEBEELKAPWIRCTHSENYSSDPM-- 427
Db 735 KNTSKSEKESENAPEDGKGNKHKRKRKKESESGLFNLPELETSRK-----SRPTGV 790
QY 428 -DQVGDSTVVGTSR-----LRDLYDKFEE--LGSROKAKAAPRPMWPKTLDLE 478
Db 791 EENKTDLSLVLPSRDATPVRD--EPMDESTITFKSVSEKDKRER---DKPAKGDKTKR 845
QY 479 SSSSECESEDSTCSSSDSE-----VPDVIAETRKKAHPRL 518
Db 846 KNDGSVSKKENIVKPAKGPQEKVDGVRDLDLNLQLKKPKRLRL 893

RESULT 13

T06291
extein homolog T9E8.80 - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C.Accession: T06291
R.Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Meyer, K.F.X.; Schuel
submitted to the Protein Sequence Database, April 1999
A.Reference number: Z15588
A.Molecule type: DNA
A.Accession: T06291
A.Residues: 1-760 <BEV>
A.Cross-references: UNIPROT:Q9TOK5; UNIPARC:UPI000009PB81; EMBL:AL049608
A.Experimental source: cultivar Columbia; BAC clone T9E8
C.Gene(s):
A.Map position: 4
A.Note: T9E8.80

Query Match 3.7%; Score 275; DB 2; Length 760;
Best Local Similarity 35.1%; Pred. No. 4.7e-07;
Matches 88; Conservative 15; Mismatches 96; Indels 52; Gaps 12;
QY 14 PGRGCPRGHGHGARSAPSPONTL-----RLHPQOPVQYQYEPSSAS----- 60
Db 361 PGRPARORSQCQCAFSLRP---PVNGSFSGGSVSP-RPVVTPLRPSSLSPRPAPRI 416
QY 61 ----TTFNSSPAPNLPRLPRPD--VPRPMPESAQGLRPCCIRPFPVNHQNHPPVP 115
Db 417 FSPRLTLTSPRPS--PRPPVSPRPPRPVYSPRPPRPVYSPRPPRPVYSPRPPRP 474
QY 116 CRRPM---PRPMCPNRPVPGARPGCGTFRFMMPRSMHRRPVY-----MQ 162
Db 475 PRPPVSPRPPSPRPPRPVYSPRPP---PRPPRPVYSPRPPRPVYSPRPPSPRPA 530
QY 163 QVNYQVPGYSHNPPPSFNSFQNNP--SSFLPSANSSSPHF--RHLPPYPL----- 212
Db 531 PLYCTPRPPPPRPPSPRPPQSPRPPRYYSRPPPHSSSPRPPSPRPPSPRPPYI 590
QY 213 -PKAPSERSP 222
Db 591 PRPPRPVSP 601

RESULT 14

T48946

hypothetical protein T15B3.60 - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C.Accession: T48946
R.Jordan, N.; Bangert, S.; Wiedelmann, R.; Voess, H.; Unseld, M.; Mewes, H.W.; Rudd, S.; I
submitted to the Protein Sequence Database, April 2000
A.Reference number: Z25009
A.Accession: T48946
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1531 <JOR>
A.Cross-references: UNIPROT:Q9LKW7; UNIPARC:UPI000009DAR8; EMBL:AL163975; GSPDB:GN0061;
A.Experimental source: cultivar Columbia; BAC clone T15B3
A.Gene: ATSP:T15B3.60
A.Map position: 3
A.Introns: 45/2; 100/3; 138/3; 183/3; 227/3; 278/3; 322/1; 357/2; 421/3; 475/2; 584/1; 67

Query Match 3.7%; Score 274.5; DB 2; Length 1531;
Best Local Similarity 21.3%; Pred. No. 1.2e-06;
Matches 247; Conservative 142; Mismatches 420; Indels 351; Gaps 58;
QY 372 WSDNOSGDKNYTSIKEKEPETWPKNEEBEELKAPWIRCTHSENYSSDPMQVG 431
Db 455 YSQQKAKQSKSKIMPL--ERGNPKQRDHLDMRREVLIDPBARLXSCPPRVKNG 512
QY 432 DST-VVGTSRRLDLYDKFEELGSRQEKAKAAPWE---DP--KTKLDED-LBSSSES 483
Db 513 HGVEIGSVIVDSNITVSEEAASQTQMSD--PSRNEQLPPCKLRDNNLQSNKGE 569
QY 484 ECESEDSTCSSSDSEVPDVIAETRKKAHNRDLDEL---WYNDPQGMNGP----- 534
Db 570 KVASSKSKSSSSSAASK-----KRKLHETTCANALSGTW---GSDIGATFOAYK 617
QY 535 ---LKKCSAK-----ARRTG-IRRSIYGEBAIK---PCRPM----- 564
Db 618 PFFCCNIGSEYVSFSLLESTLADGVKEMDVLVRLVYASVSPCCQILSLGSELYK 677
QY 565 -----TNNAGRLFHYRITVSPPTNFLTDRPTVIEYDHEVYFEGSFMAHAPLTI 615
Db 678 AKYQOQFFNGMFGTLFVQSKSQGTKEBLLQ-----TDTSILMHPAFMELLPVETN 730
QY 616 PLCKVIRRNIDVTIHFIEEMPEMPCVKGLEFLSLFPLRD-LLELYDNMLKORLPEDSP 674
Db 731 DLAS--SATIDWS-----AINSCASIVE---FLKNSLILDLRD-----SDGNQ 768
QY 675 CCPRFHMPFRVRLPDGKEVLSMHQILLYLRCSKALVPEEELANMLQWELEWQKYA 734
Db 769 C-----NTSSQGEVL-----LDDGMEEETNLHFANASSDKNS 800
QY 735 EECKGMIVTNPCTKPSVRI--DQLDRQFNDVITFPIIVHFGIRPAQLSYAGDPQYK 792
Db 801 LBEVLVIAIHTRISYIYEAVSDSSAMSPFEVD-----ASSGATVAYEYN 846
QY 793 LMKSY-VKLRIHLANS--KYKQTDKQ---LAQREALQKIRQK--NMRREVYELSS 844
Db 847 --KKTGIVLDA--PQOPLMKLKQSHHANLVDENEEMVVKTEPKAGNKKRPNI----- 898
QY 845 QGFMTGTGRSDVCGAMMLPVLTHIRYHQCIMHLDKLIGTFODRCLQLAMTHPS--H 902
Db 899 -----HAHLPEL-----LARD-----VPAVALKSIYLLPSVNH 928
QY 903 HLNFGNPDHARNSLNSCGIRQPKYGDGRKVNHHMKKGINTLINIMSRIGQDDPTPSRI 962
Db 929 RLSESLMASQLREEL-DCSIDNFSISTSI-----LEAVTTL-----TCPRS 969
QY 963 NINERLEFLGDAVVEFLTSVHLYYLPSSLSEGLATYRAIYQNGHLNL--AKKE--- 1017
Db 970 FSEMRLEELGDSVLTKVYASCHLFLKYPDKDEQOLSROROSIISSNLAHLTTSRKLQGYI 1029
QY 1018 ----LDPMFLVAHG-----PDLG-----RESLRH----- 1038
Db 1030 RNGAFEPRRWTAPGQSLFPVPCCKGIDTRVPLDPKFTENMTIKIGKSCMGHKNVVS 1089

QY	1039	---AMANCBEALIGAVYLEGSLSEEAQOLFGRLLFTN---DPDIREVLANYPRLHLOQEBNTD	1094
Db	1090	KSVSDCBEMALIGATVYSGSLSHAMMKWIGLIDVDPBNLVVEAINV---SLRCYIKRED	1147
QY	1095	ROLIETSPVLQKLTFFEEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHN---QMBFLGD	1151
Db	1148	-----ELIETLRKIQHESFAKFLLEKAIT-----HSLRESYSVERLEFJGD	1189
QY	1152	SIMQVATETVLFTHFPRDHGCHLTLSSLSLVNRTQAKVAE-----ELGM	1196
Db	1190	SVADFLTLRHLFNTYEQTGPREMTDLRSACANNINPQVAKKNNLTHLORCATVLETOI	1249
QY	1197	QEVAIT---NDKTKR---PVGRLRTKLADLLESFATAALYTDBOLEYVTHFMNVCFEPRLKE	1251
Db	1250	NDVILMSQKQDETCRSIPSTIOGPKALGDVVESTIGALLIDTRDLDDQVMR---VEPILSP	1307
QY	1252	FIINQDWN-DPKSQLOQCCULT---RTEGKERDIPRYKTLQ-----TVGSPHARTYT	1299
Db	1308	LVTTPDKLQLPYRELNELCSLGYFPRVKCSNDGVAKQATIQLDDBVLLTGDSSEQT--	1365
QY	1300	VANYFKERIGCGCGSPSIOQAEMGAMDALEKYNFQMAHQKRTIGKY---RQELKEM	1355
Db	1366	-----NKLALGQASHLLTLQJ-----EKRNISRTKSIGNQSSMDV	1401
QY	1356	RWEREHOERE--PDETEDIK	1373
Db	1402	NLACNHSDRETLTSSETTEIQ	1421

Db 197 GQQT- QEPPLW-----LPPPPFGD-ETPPVPSLPPPLDEPPP-----MPITWLPPPPVPA 246

QY 218 ERSPPRLKG-----YDDRRHRDHSHGGERHSLSRRRGRSPDDRQDS 263

Db 247 QISSAALFQIPLPVITTEALENPNHSRRHD-----ENKKGDLDRNR-RVYSRRRSRS 299

QY 264 R 264

Db 300 R 300

Search completed: February 3, 2006, 19:42:06
Job time : 57 secs

```

RESULT 15
G86292
hypothetical protein F7H2_17 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86292
R/Theologian: A.; Eckert, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hutzar, L.
Nucleotide 408, 816-820, 2000
A/Author: Hunter, Y.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Author: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: G86292
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1006 <STO>
C/Cross-references: UNIPROT:Q9LMQ1; UNIPARC:UPI00000A2407; GB:AB005172; NID:g8927662; P
C/Genetic:
;/Map position: 1

```

Query Match 32.6% Score 273.5 DB 2 Length 1006;
 Best Local Similarity 3.9% Pred. No. 8e-07;
 Matches 99; Conservative 17; Mismatches 98; Indels 87; Gaps 19;

```

QY      3 QGNTCHAMSHPRAGCPRGRG-----GHGAPASAPSRPQURLRLHNGQR 47
Db      48 QSNRGRGSS-----GSSRLNRGRSCDPLYOYLPTCGHWPPPTPS--PENPLF--PPQP 98

QY      48 PVOYQYEPSPASPTTFNSPAINFLLPRRPDVP--FPPMPSPAGPLRPCEIRDPFNHQ 106
Db      99 P-----RPPRPFR-----DRPRSPRLPPLVLVSPPLPLHPR--SPCEPPL 138

QY      107 MRHPFVYPCCPMPMPCCNNPVPVGAPGQGTFP-----PMMPSPSMHPRPVPM 160
Db      139 MPSPPLVPS--PPLPPSPSLVPSPP--PPSPPPFFPSPRPVVLVPLVPSPPPLPG 196

QY      161 PQQVYVYPPGYSHHNPPEPSGFNSFQNNPSFLPSANSSSPHFRHLDP--YLPKAPS 217

```

THIS PAGE BLANK (USPTO)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:32:44 ; Search time 287 Seconds
(without alignments)
3377.686 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMQNTCHRMSPHGRGCRP.....MRMRSHORREPDATEDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7471	99.6	1374	1 RNC_HUMAN	Q9nrt4 homo sapien
2	7160	95.5	1373	2 Q5H2J0_MOUSE	Q5h2j0 mus musculu
3	6380	85.1	1231	2 Q6PAJ6_MOUSE	Q6paj6 mus musculu
4	6344.5	84.6	1336	2 Q5ZIR3_CHICK	Q5zir3 gallus gall
5	5217	69.6	1021	2 Q80Z69_MOUSE	Q80z69 mus musculu
6	4982	66.4	942	2 Q6PF88_MOUSE	Q6pf88 mus musculu
7	4121	54.9	780	2 Q7TW18_MOUSE	Q7tw18 mus musculu
8	3767	50.2	1041	2 Q4SKM1_TETNG	Q4skm1 tetraodon n
9	3696	49.3	705	2 Q80UJ5_MOUSE	Q80uj5 mus musculu
10	2963	39.5	1327	2 Q9XN5_DROME	Q9xyn5 drosoephila
11	2881.5	38.4	979	2 Q5TKR8_ANOGA	Q5tkr8 anopheles g
12	2825	37.7	541	2 Q8CJ74_MOUSE	Q8cj74 mus musculu
13	2811.5	37.5	1071	2 Q960Y4_DROME	Q960y4 drosoephila
14	2683.5	35.8	836	2 Q7PNE9_ANOGA	Q7pne9 anopheles g
15	2298	30.6	432	2 Q59FF9_HUMAN	Q59ff9 homo sapien
16	1480	19.7	288	2 Q67B98_BAT	Q67b98 ratcus norv
17	1473.5	19.6	1061	2 Q61XX9_CABER	Q61xx9 caenorhabdi
18	1448	19.3	1086	1 RNC_CABER	Q01336 caenorhabdi
19	1154.5	15.4	290	2 Q4SKM3_TETNG	Q4skm3 tetraodon n
20	1138.5	15.2	412	2 Q569M4_XENLA	Q569m4 xenopus lae
21	711	9.5	134	2 Q9CGM2_MOUSE	Q9cgm2 mus musculu
22	344.5	4.6	220	2 Q82ZG1_ENTRA	Q82zg1 enterococcu
23	340.5	4.5	1150	2 Q81M66_DROME	Q81m66 drosoephila
24	340	4.5	846	2 Q8MRP6_DROME	Q8mrp6 drosoephila
25	340	4.5	1109	2 Q9VAV4_DROME	Q9vav4 drosoephila
26	339	4.5	1389	2 Q5SFS1_DICDI	Q5sfs1 dictyosella
27	331	4.4	1838	2 Q6BCJ9_TETNG	Q6bcj9 tetrahymena
28	329.5	4.4	2869	2 Q60PD2_CABER	Q60pd2 caenorhabdi
29	328	4.4	1022	2 Q4SFX9_TETNG	Q4sfx9 tetraodon n
30	325	4.3	551	2 Q8BX18_MOUSE	Q8bx18 mus musculu
31	321	4.3	249	2 Q65J05_BACUD	Q65j05 bacillus

32	321	4.3	551	2 Q8BX86_MOUSE	Q8bx86 mus musculu
33	321	4.3	551	2 Q6NVF9_MOUSE	Q6nvf9 mus musculu
34	319	4.3	825	2 Q7S9H3_NEUCR	Q7s9h3 neuropeptid
35	317.5	4.2	552	2 Q5NVH8_PONPY	Q5nvh8 ponopysoma
36	316	4.2	551	2 Q5TBS1_HUMAN	Q5tbs1 homo sapien
37	316	4.2	551	2 Q16630_HUMAN	Q16630 homo sapien
38	315	4.2	249	1 RNC_BACSU	P51833 bacillus su
39	311.5	4.2	528	2 Q4SQ65_TETNG	Q4sq65 tetraodon n
40	310	4.1	548	2 Q6DWM4_XENLA	Q6dwm4 xenopus lae
41	310	4.1	551	2 Q5ZL34_CHICK	Q5zl34 gallus gall
42	309.5	4.1	1604	2 Q7X014_ORYSA	Q7x014 oryza sativ
43	308.5	4.1	229	1 RNC_BACCR	Q81978 bacillus ce
44	307.5	4.1	225	1 RNC_LISMO	Q81971 bacillus mo
45	307.5	4.1	245	1 RNC_BACAN	Q81w18 bacillus an

ALIGNMENTS

RESULT 1
ID RNC_HUMAN STANDARD: PRT: 1374 AA.
AC Q9NRR4; Q7Z5V2; Q86YH0; Q9NM73; Q9YZV9; Q9Y4T0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III) (Drosna) (p241).
GN Name=RNASE3; Synonyms=RN3;
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=20538440; PubMed=10948199; DOI=10.1074/jbc.M005494200;
RA Wu H., Xu H., Miraglia L.J., Crooke S.T.;
RT "Human RNase III is a 160-kDa protein involved in preribosomal RNA
processing";
RL J. Biol. Chem. 275:36957-36965(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 166-613 (ISOFORM 2).
RC TISSUE=Colon;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Matita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagatani K., Murakami K., Yasuda T., Iwayanagi T., Magatsuna M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamaaki M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
RA Iihara S., Ono Y., Takiguchi S., Watanabe S., Yosiida M., Hotura T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togashi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Mutsaers H. K., Yuuki H., Oshima A., Saeki N., Aoeuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nishigawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe K., Kumagai A., Itakura S., Fukumizu Y.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukumizu Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami H., Fujiiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohnori Y.,

RA Kikwabati A., Hkikiti T., Koblate S., Inogaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Motoringa M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuo Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Inogaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RN NCULEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3), AND NCULEOTIDE
RP SEQUENCE [LARGE SCALE MRNA] OF 653-1374 (ISOFORM 1).
RC
RC TS3NUC=Cervix, and Skin; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Groner L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hailton E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimson J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzyzaniak M.I., Skalska U., Smalins D.E.,
RA Schmeich A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RN NCULEOTIDE SEQUENCE OF 706-1374.
RC TS3UC=Aorta;
RA Wei Y.J., Ding J.F., Xiong H., Zhou Y., Liew C.C.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RN PubMed=14508493; DOI=10.1038/nature01957;
RX Lee Y., Ahn C., Han J., Choi H., Kim Y., Yim J., Lee J., Provost P.,
RA Radmark O., Kim S., Kim Y.N.;
RT "The nuclear RNAse III Drosha initiates microRNA processing.";
RL Nature 425:415-419(2003).
CC -1- FUNCTION: Executes the initial step of microRNA (miRNA) processing
CC in the nucleus, that is cleavage of pri-miRNA to release pre-
CC miRNA. Involved in pre-RNA processing. Cleaves double-strand RNA
CC and does not cleave single-strand RNA.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- SUBUNIT: Interacts with Spl.
CC -1- INTERACTION:
CC QMRY05:DCRR; NBExp=2; Intact=EBI-528367, EBI-528411;
CC -1- SUBCELLULAR LOCATION: Nuclear. A fraction is translocated to the
CC nucleolus during the S phase of the cell cycle.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9NRR4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NRR4-2; Sequence=VSP_005777;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q9NRR4-3; Sequence=VSP_012450, VSP_012451, VSP_012452,
CC VSP_012453;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
CC -1- SIMILARITY: Contains 2 RNase III domains.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to a

[illegible]

```
QY 241 GERRHSLDRRRGRSPDRRRRQDSRYSDYDRGRTPSRHRSEYERSRERERHRHRDRNRSS 300
DB 241 GERHRLDRRRGRSPDRRRRQDSRYSDYDRGRTPSRHRSEYERSRERERHRHRDRNRSS 300
QY 301 PSLEERYKKEKYSKSGSYGLSVYEPAGCTPELPGEIIKKTDSWAPPLEIVNHRSPSRK 360
DB 301 PSLEERYKKEKYSKSGSYGLSVYEPAGCTPELPGEIIKKTDSWAPPLEIVNHRSPSRK 360
QY 361 KKAARWEEEDKDRMSDNOSSGDKKNTYSIKEKEPEETPDKNKEESEEELKRWIRCHSEN 420
DB 361 KKAARWEEEDKDRMSDNOSSGDKKNTYSIKEKEPEETPDKNKEESEEELKRWIRCHSEN 420
QY 421 YSSSDMDVQDSTVYVTSRLRDLYDKFBEELGSRQEKAKAARPWEPKTKLDEDLSS 480
DB 421 YSSSDMDVQDSTVYVTSRLRDLYDKFBEELGSRQEKAKAARPWEPKTKLDEDLSS 480
QY 481 SSESCESDSDSTCSSSSDSEVPVIAIKKKKAPRLHDELMYNDPGQWMDGPLCKCSA 540
DB 481 SSESCESDSDSTCSSSSDSEVPVIAIKKKKAPRLHDELMYNDPGQWMDGPLCKCSA 540
QY 541 KARRGTIRHSIYGEBAIKPCRPMTNAGRLPHYRTTVSPNMLDRPTVIEYDHEXI 600
DB 541 KARRGTIRHSIYGEBAIKPCRPMTNAGRLPHYRTTVSPNMLDRPTVIEYDHEXI 600
QY 601 PEGFSMFAAPLTNIPLCVKIRFNIDYTHIFIEBMPENFCVGLFSLFLPRDILELY 660
DB 601 PEGFSMFAAPLTNIPLCVKIRFNIDYTHIFIEBMPENFCVGLFSLFLPRDILELY 660
QY 661 DNNLKGPLEFDSPPCCPRFHPMRPVRLPDGKEVLSMHQILLYLRCSKALVPEEIA 720
DB 661 DNNLKGPLEFDSPPCCPRFHPMRPVRLPDGKEVLSMHQILLYLRCSKALVPEEIA 720
QY 721 NMLQWELEMOKAYAECKMIVTNPCTKSSVNIIDLDREOFDPDVTTPPIVHFIRPA 780
DB 721 NMLQWELEMOKAYAECKMIVTNPCTKSSVNIIDLDREOFDPDVTTPPIVHFIRPA 780
QY 781 QLSYAGDPQYQKLMKSYVLYLHLLANSPKVKOTDKOKLQREBALKIKOKMTREEVY 840
DB 781 QLSYAGDPQYQKLMKSYVLYLHLLANSPKVKOTDKOKLQREBALKIKOKMTREEVY 840
QY 841 ELSSOGFMKTGIRSDVCOHAMMLPVLTTHIIRYHQCIMHLDKLGYTFQDRCLQLAMTHP 900
DB 841 ELSSOGFMKTGIRSDVCOHAMMLPVLTTHIIRYHQCIMHLDKLGYTFQDRCLQLAMTHP 900
QY 901 SHHLLNGMNPDAHARNLSNCGIRQPKYGRKYNHMRKKGINTLINTMSRLGQDDPTPS 960
DB 901 SHHLLNGMNPDAHARNLSNCGIRQPKYGRKYNHMRKKGINTLINTMSRLGQDDPTPS 960
QY 961 RINNHRLREFLGDAVVEFLTSVHLYLPSLEEGGLATYRTAIVONQHILAMLAKELEDP 1020
DB 961 RINNHRLREFLGDAVVEFLTSVHLYLPSLEEGGLATYRTAIVONQHILAMLAKELEDP 1020
QY 1021 FMUYAHGPDLCESBDRHAMANCFEALIGAVYLEGSLBEAKOLFGLFLFNDPLREVMIN 1080
DB 1021 FMUYAHGPDLCESBDRHAMANCFEALIGAVYLEGSLBEAKOLFGLFLFNDPLREVMIN 1080
QY 1081 YPLHPLQLOEPNTDROLITSPYLOKLTETEPEEAGIYFTHVRLAAAFLLRTYGFNHLTL 1140
DB 1081 YPLHPLQLOEPNTDROLITSPYLOKLTETEPEEAGIYFTHVRLAAAFLLRTYGFNHLTL 1140
QY 1141 GHNORNEFLGDSIMOLVATEYLPFHPDHHEGLTLNRSIVNNRQOAVAEELGQCEYA 1200
DB 1141 GHNORNEFLGDSIMOLVATEYLPFHPDHHEGLTLNRSIVNNRQOAVAEELGQCEYA 1200
QY 1201 ITNDTKRBPVGLRTKTLADLBSFIALYTDKOLEVHTFMNVCFEPLKEFLINDAMD 1260
DB 1201 ITNDTKRBPVGLRTKTLADLBSFIALYTDKOLEVHTFMNVCFEPLKEFLINDAMD 1260
QY 1261 PXSOLQOCCILTRTEKEPEDIPLYKTLQVYVGSNANTYVAVYFKKERIGCGKGPSIQQA 1320
DB 1261 PXSOLQOCCILTRTEKEPEDIPLYKTLQVYVGSNANTYVAVYFKKERIGCGKGPSIQQA 1320
```

```
QY 1321 EKGAMDALEKTNFQMAHQKRFGRKGYOELKEMWEREHOEREDFEDIDIK 1374
DB 1321 EKGAMDALEKTNFQMAHQKRFGRKGYOELKEMWEREHOEREDFEDIDIK 1374

RESULT 2
05H2J0 MOUSE PRELIMINARY; PRT; 1373 AA.
AC 05H2J0;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Rnasen protein.
GN Name=Etohl2; Synonyms=Rnasen;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo W.F., Casavant J.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullady S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney U., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC088999; AAH88999.1; -; mRNA.
DR Ensemble; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etohl2.
DR MGI; MGI:1261425; Rnasen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; Ds_RNA_bd.
DR InterPro; IPR000959; Rnae_III.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00358; Rbounclase_3; 2.
DR SMART; SM00355; RIBOC; 2.
DR PROSITE; PSS0137; DS_RBD; 1.
DR PROSITE; PSS0137; RNASE_3_1; 2.
DR PROSITE; PSS0142; RNASE_3_2; 2.
SQ SEQUENCE 1373 AA; 158828 MW; EDIDEEZ2FFD89A6B CRC64;

Query Match 95.5%; Score 7160; DB 2; Length 1373;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1310; Conservative 32; Mismatches 30; Indels 2; Gaps 2;
```

Db 1 MCGNTCHRSYHPRGCGRGGHGARPSAPAFRPNLRLHQPDPAPQYQTEPPSAPS 60
Qy 62 TGSNBPAPFLPRPDPFVPPPMPSAAGPLPCCIRPPFNHQMHRPVPVPCPPMP 121
Db 61 SSSNQAAPFMPBRPFPVYPPAPASAGPLPCCVVRPPYNNQHRRFPVPPCPRMP 120
Qy 122 PMPCCNPPVPGAPPGQCTFPPMPDPSPMPHPPPPVMPQQVNYQYPPGYSHNPPPS 181
Db 121 PMPCCNPPVPGAPPGQCTFPPMPDPSPMPHPPPPVMPQQVNYQYPPGYSHNPPPS 179
Qy 182 PMSFONPSSFLPSANSSSPHPRHLPPYPLPKAPERSPERLKHYYDHRHDSHGK 241
Db 180 PMSYQNNSSFPSPASSTPHPRHLPPYSLPQAQNNRSPELKHYYDHRHDSHGK 239
Qy 242 EHRHSLDRERGRSPDRRQDSKRSYDVRGRT-PSRRHSYERSRERERHRHRDRRS 300
Db 240 EHRHSLDRERGRSPDRRQDSKRSYDVRGRT-PSRRHSYERSRERERHRHRDRRS 299
Qy 301 PSLERSYKKEKYSRSGSYGLSVPEPAGCTPELPGELIKNTDSMAPPLEIVNHRSPRK 360
Db 300 PSLERSYKKEKYSRSGSYGLSVPEPAGCTPELPGELIKNTDSMAPPPENVHRSPRK 359
Qy 361 KARPEEEDRWSDNOSGKOKKYTSIKKEKEPEMMDKNEEERELPKVMIRCHSHN 420
Db 360 KARPEEEDRWSDNOSGKOKKYTSIKKEKEPEMMDKNEEERELPKVMIRCHSHN 419
Qy 421 YSSDPMDOGVSTVGTGRLRLDYKFEELGSRQEKAKAARPMPEPKTLDEDESS 480
Db 420 YSSDPMDOGVSTVGTGRLRLDYKFEELGSRQEKAKAARPMPEPKTLDEDESS 479
Qy 481 SESECESDSDSTCSSSDSEVPDVAIEIKRKAHPRLHDELMYNDPGQNDGPKCSA 540
Db 480 SESECESDSDSTCSSSDSEVPDVAIEIKRKAHPRLHDELMYNDPGQNDGPKCSA 539
Qy 541 KARPGIRSIYRGEBAIKPCRMNTNNAARLPHYRITVSPPNPLDRDTVIEYDHEXI 600
Db 540 KARPGIRSIYRGEBAIKPCRMNTNNAARLPHYRITVSPPNPLDRDTVIEYDHEXI 599
Qy 601 FCGFSMFAAPLPLNIPLCVKIRFNIDYTHIFIEEMMPENFCVGLFPLFRDILLEY 660
Db 600 FCGFSMFAAPLPLNIPLCVKIRFNIDYTHIFIEEMMPENFCVGLFPLFRDILLEY 659
Qy 661 DNMLKGPLFEDSPCCPRPHMPRPVRLPDGKEVLSNHQIILYLLRCSKALVPEEIA 720
Db 660 DNMLKGPLFEDSPCCPRPHMPRPVRLPDGKEVLSNHQIILYLLRCSKALVPEEIA 719
Qy 721 NMLQWELMOKYAECKKMIYNNPCTKSSVPIIDOLDBEQFNPDTITPIVHGIRRA 780
Db 720 NMLQWELMOKYAECKKMIYNNPCTKSSVPIIDOLDBEQFNPDTITPIVHGIRRA 779
Qy 781 QLSYAGDPQYQKLMSYVYKLRHLNLAASPKYQTDOKLAQREBALQIRQKMTREVV 840
Db 780 QLSYAGDPQYQKLMSYVYKLRHLNLAASPKYQTDOKLAQREBALQIRQKMTREVV 839
Qy 841 ELSSQGFMTGTGRSDVCOHAMMLPVLTHHRYQCMLHDKLIGYTFODRCILQMLMTP 900
Db 840 ELSSQGFMTGTGRSDVCOHAMMLPVLTHHRYQCMLHDKLIGYTFODRCILQMLMTP 899
Qy 901 SHHNGMNPDRARNLSNCGIRQPKYGRKTHMMMRKGINTLNINSRIGODDPTPS 960
Db 900 SHHNGMNPDRARNLSNCGIRQPKYGRKTHMMMRKGINTLNINSRIGODDPTPS 959
Qy 961 RINHNRLFFLGDAVVEFLTSVHLVYLLPSLEGGATRYTAIVQNHMLAKKLELP 1020
Db 960 RINHNRLFFLGDAVVEFLTSVHLVYLLPSLEGGATRYTAIVQNHMLAKKLELP 1019
Qy 1021 FMLYAHGPDLCRESDLRHAMANGFEALIGAIVLEGLLEBAKQIFGLIFNDPLREVMIN 1080
Db 1020 FMLYAHGPDLCRESDLRHAMANGFEALIGAIVLEGLLEBAKQIFGLIFNDPLREVMIN 1079
Qy 1081 YPLHPIQLOBPNTDROLIETSPVLQKLTREBAIGVIFTHVRLAARAFILRTVGFMHLL 1140
Db 1080 YPLHPIQLOBPNTDROLIETSPVLQKLTREBAIGVIFTHVRLAARAFILRTVGFMHLL 1139

Qy 1141 GHNQMEFLGDSIMQVATETYLFIHPDHHEGHTLLASSLVNNRTQAVAEELGQVEYA 1200
Db 1140 GHNQMEFLGDSIMQVATETYLFIHPDHHEGHTLLASSLVNNRTQAVAEELGQVEYA 1199
Qy 1201 ITNDKTKRPVGLRTKTLADLESFIALYTTDQLEXYVHFMVNCFFPRLKEPILINDMND 1260
Db 1200 ITNDKTKRPVGLRTKTLADLESFIALYTTDQLEXYVHFMVNCFFPRLKEPILINDMND 1259
Qy 1261 PKSQLQCCCLTLRTEGKEPDIPLVYKTLQTVGSHARTYVAVYFKGERIGCKGSPISQQA 1320
Db 1260 PKSQLQCCCLTLRTEGKEPDIPLVYKTLQTVGSHARTYVAVYFKGERIGCKGSPISQQA 1319
Qy 1321 EWGAAMDALKEKNFQMAHQKRFGRKTRQELKEMRWEREHQEREPDEIEDIK 1374
Db 1320 EWGAAMDALKEKNFQMAHQKRFGRKTRQELKEMRWEREHQEREPDEIEDIK 1373

RESULT 3
Q6PAU6_MOUSE
ID Q6PAU6_MOUSE PRELIMINARY; PRT; 1231 AA.
AC Q6PAU6;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Name=Etch12; Synonym=RNaseen;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=24238857; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schler G.D.,
RA Altchul S.F., Zeeberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.U., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (OCT-2003) to the EMBL/Genbank/DBD databases.
DR EMBL; BC060265; AAH60265.1; -; mRNA.
DR Ensembl; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etch12.
DR MGI; MGI:1261425; RNaseen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; Ds_RNA_bd.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dsm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.

DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS00142; RNASE_3_2; 2.
FT NON_TER 1 1
SQ SEQUENCE 1231 AA; 143573 MW; 3679FA9D87800151 CRC64;

Query Match 85.1%; Score 6380; DB 2; Length 1231;
Best Local Similarity 96.1%; Pred. No. 2.9e-303;
Matches 1164; Conservative 22; Mismatches 24; Indels 2; Gaps 2;

QY 144 FMWPEPMHPPEPPVWQVNYQYPPGYSHNFPPEPSFNSFONNPSFLPSANSSSPH 203
DB 1 FMWPEPMHPPEPPVWQVNYQYPPGYSHNFPPEPSFNSFONNPSFLPSANSSSPH 59
QY 204 FRLPLPYPLPKASERKSPERLKYDDHRRDHSNGRGRRLDRLRRERKSRDRRSDS 263
DB 60 FRLPLPYSLPKAONERRSPERLKYDDHRRDHSNGRGRRLDRLRRERKSRDRRSDS 119
QY 264 RYASDIDRGRT-PSRRSYRERERERERHRRDNRSPSLERSYKKEYSGRSGSLV 322
DB 120 RYASDIDRGRTPPRRHRSYRERERERHRRDNRSPSLERSYKKEYSGRSGSLV 179
QY 323 VEPACCTPELPGELIKNTDSMAPLEIVNHRSPSEKRRARWEBEKDRWSDQSGKDK 362
DB 180 AEPACCTPELPGELIKNTDSMAPLEIVNHRSPSEKRRARWEBEKDRWSDQSGKDK 239
QY 383 NNTSIEKEPEEETMPDKNEEBEELKPVWIRCTHSBNYSSDPNQVGDSTVYGRSLR 442
DB 240 NNTSIEKEPEEETMPDKNEEBEELKPVWIRCTHSBNYSSDPNQVGDSTVYGRSLR 299
QY 443 DLYDKEEELSGNQEAKAAPRWPBPCKTKLDEDLSSSECESDSDSTCSSSDSEVF 502
DB 300 DLYDKEEELSGNQEAKAAPRWPBPCKTKLDEDLSSSECESDSDSTCSSSDSEVF 359
QY 503 DVAIAIKRKKAHDRHDELWYNDPGOMDGPLCKCSAKARRTIGHSIYPGELAKPCR 562
DB 360 DVAIAIKRKKAHDRHDELWYNDPGOMDGPLCKCSAKARRTIGHSIYPGELAKPCR 419
QY 563 PMTNNAGRLEFHYRITVSPPTNFLTDRPTVIEYDHEVIFEGSGMFAHAPLTNIPCKVIR 622
DB 420 PMTNNAGRLEFHYRITVSPPTNFLTDRPTVIEYDHEVIFEGSGMFAHAPLTNIPCKVIR 479
QY 623 FNIDYIHFEEMMPNFCVKGLEFSLFRDILEYDMNKGPLFEDSPCCPFHFM 662
DB 480 FNIDYIHFEEMMPNFCVKGLEFSLFRDILEYDMNKGPLFEDSPCCPFHFM 539
QY 683 PRFVRPLPDGKKEVLNMQILLYLRCSKALVPEEBEIANMLQWEELEWQYAECKGMIV 742
DB 540 PRFVRPLPDGKKEVLNMQILLYLRCSKALVPEEBEIANMLQWEELEWQYAECKGMIV 599
QY 743 TNPGRKSSVRIDQDREQFNPDVITFPIIVHFGIRPAQLSVAGDPQYQKLKSYVKLRH 802
DB 600 TNPGRKSSVRIDQDREQFNPDVITFPIIVHFGIRPAQLSVAGDPQYQKLKSYVKLRH 659
QY 803 LLANSKXVOTDKQKLAOREEALOKIRKQKTMREVEVLSGCFKGTGRSVCOHAMM 862
DB 660 LLANSKXVOTDKQKLAOREEALOKIRKQKTMREVEVLSGCFKGTGRSVCOHAMM 719
QY 863 LPLVLTNHRHOCIMHLDKLGTYTPODRCLQLAMTHPSHNLNFGNPDHARSLSNCGI 922
DB 720 LPLVLTNHRHOCIMHLDKLGTYTPODRCLQLAMTHPSHNLNFGNPDHARSLSNCGI 779
QY 923 ROPKYGDRKVNHHMKKGINVLINIMSRIGODPTPSRINHNRELEPLGDAVVEFLTSV 982
DB 780 ROPKYGDRKVNHHMKKGINVLINIMSRIGODPTPSRINHNRELEPLGDAVVEFLTSV 839
QY 983 HLYULFPESEEGGLAYRTAIVONOHILAMAKKLEIDRFLVHAGDPLCESDRLRAMAN 1042
DB 840 HLYULFPESEEGGLAYRTAIVONOHILAMAKKLEIDRFLVHAGDPLCESDRLRAMAN 899
QY 1043 CFEALIGAVYLEGSLTEAKQLFGRLTFNDPDLREVWLVNPLHPLQOEPTDROLIETSP 1102

DB 900 CFEALIGAVYLEGSLTEAKQLFGRLTFNDPDLREVWLVNPLHPLQOEPTDROLIETSP 959
QY 1103 VLOKTEPEEALGVITHTHRLILARAFLLRTVGFNNHILGHONMEELGDSIMOLVATEYL 1162
DB 960 VLOKTEPEEALGVITHTHRLILARAFLLRTVGFNNHILGHONMEELGDSIMOLVATEYL 1019
QY 1163 FIHPDHHGHLTLRSLSVNNRTQAKVAEELQMOEYATINDTKRPVGLRTKTLADLLE 1222
DB 1020 FIHPDHHGHLTLRSLSVNNRTQAKVAEELQMOEYATINDTKRPVGLRTKTLADLLE 1079
QY 1223 SFIAALYTDKLEAYTFNANVCFFPLKEFILNQDNDPKSQLOCCCLTRTEGKEPDIP 1282
DB 1080 SFIAALYTDKLEAYTFNANVCFFPLKEFILNQDNDPKSQLOCCCLTRTEGKEPDIP 1139
QY 1283 LYKTLQTVGSPHARTVAVYFKGERIGCGKPSIQOALEGAMMDLLEKYNFQMAHOKR 1342
DB 1140 LYKTLQTVGSPHARTVAVYFKGERIGCGKPSIQOALEGAMMDLLEKYNFQMAHOKR 1199
QY 1343 FIERKTYROELKEMWEREHOEREPDETDIKK 1374
DB 1200 FIERKTYROELKEMWEREHOEREPDETDIKK 1231

RESULT 4
Q5ZIR3 CHICK PRELIMINARY; PRT; 1336 AA.
ID Q5ZIR3 CHICK PRELIMINARY; PRT; 1336 AA.
AC Q5ZIR3
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.24a21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blasodatski A., Kostovska D., Kotler M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis";
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ720721; CAG32380.1; -; mRNA.
DR GO; GO:0005634; Cnucleus; IEA.
DR GO; GO:0003725; P.double-stranded RNA binding; IEA.
DR GO; GO:0004519; Pendonuclease activity; IEA.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR GO; GO:0004525; F.ribonuclease III activity; IEA.
DR GO; GO:0006396; P.RNA processing; IEA.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PR00035; dsrm; 1.
DR Pfam; PR00036; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS00142; RNASE_3_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 1336 AA; 154836 MW; 50D9A9A9C53DE0CF CRC64;

Query Match 84.6%; Score 6344.5; DB 2; Length 1336;
Best Local Similarity 85.3%; Pred. No. 1.7e-301;
Matches 1171; Conservative 76; Mismatches 81; Indels 45; Gaps 5;

QY 10 MSFHPGRCPRGCGHGAPSPASRPPQVRLHLPQDPQVQYQYEPSPASPTTFSPAP 69
DB 1 MSFHPGRCPRGCGHGAPSPASRPPQVRLHLPQDPQVQYQYEPSPASPTTFSPAP 60

```

QY 70 NPLPRPDPVPPPPPPPSAOGCLRPCCPIRRPPRNQMRPPRVPPCSPMPMPMPCPN 129
  :::::::::::::::::::::
Db 61 SVPMPRPDPVPPPPPPSTQNPINOCMPMRPPPNHQMSFPVPCCFPAPBPVAPBP 120
QY 130 PVPVGPAPPGQGTFFPMMPRPMRPPPPVPMQVUNYQ--YPPGYHNHPPPSPSPN 187
  :::::::::::::::::::::
Db 121 ACPVATATQGTFFPMMPRPPVAVHPPPPPVPIPOQVUNYQVSTGVSQGSFPNPNFSYQH 180
QY 188 NSSFLPSANSS--SPHFRHLPPYLPKAPSERASPERLKYDHRHRHSHGGERAR 245
  :::::::::::::::::::::
Db 181 NSSSPGSSSTTSSSSSHFRHQYQGEKSGSDRBRPERSKHYDHRHRHSHGDRAR 240
QY 246 SI--DRRERGSPPDRROD-SKYSRDYRGRTSPSHRSYERSRERERHRHRDRRSP 301
  :::::::::::::::::::::
Db 241 SNHGDRDRGRSGSPDRRRQESSRHRTYDRGRMSPHRSYERSRERDRHRHRDRSP 300
QY 302 SIERSYKKEKYSGRSGYLVVPERAGCTPELGEIIXNTDSAPPLELVNHRSPREK 361
  :::::::::::::::::::::
Db 301 SPDRSYKKQYKAG-----SRSPREK 323
QY 362 RARWEERKRWSDNOSSGDKNYTTSIKEKPEETMPDKNEEELKPEVIRCTSENY 421
  :::::::::::::::::::::
Db 324 RRRWEERERMSQSQANKEKNTSIXDKESSENAEKDEDELLKVVWRCHTSEY 383
QY 422 YSSDPMDQVGDSTVVGTSRLRDLXKFEELSGRQEKAKAAPRPMWPKTKLDEDESSS 481
  :::::::::::::::::::::
Db 384 YSNDPMDQVGDSTVVGTSKLRLDYKFEELSGRQAKAKAAPRPMWPKTKLDEDESSS 443
QY 482 BECESEDESTGSSSDSEVPVIAIKRKKAPRDLHMLWNDGQNDGPKCKCSAK 541
  :::::::::::::::::::::
Db 444 BECESDSDSSCGSSSDSEVPVIAIKRKKAPRDLHMLWNDGQNDGPKCKCSAK 503
QY 542 AARTGTRHSIYGEZAIKPCRPMTNAGRLFHYRIIVSPPTNPLDRPVYIEVDHEIYF 601
  :::::::::::::::::::::
Db 504 AARTGTRHGIYGEZAIKPCRPMTNAGRLHYRIIVSPPTNPLDRPVYIEVDHEIYF 563
QY 602 EGSFMAHAPLNIPLCKVIRNIDYTIHFIEBMBENFCVKGLELFSFLPRDILIELYD 661
  :::::::::::::::::::::
Db 564 EGSFMSHAPLNIPLCKVIRNIDYTIHFIEBMBENFCVKGLELFSFLPRDILIELYD 623
QY 662 NMLKGLFEDSPCCRRPHMPRVFPLDGGKEVLSMOIILYLKCSKALVPEEIN 721
  :::::::::::::::::::::
Db 624 MHLKGSFVNDVSVCRRFHPMFVFLPDGGKEVLSMOIILYLKCSKALVPEEIN 683
QY 722 MLOMEELEROKYAECKGMIVTNPGTKPSSVRIDQDREQFNDVITPFIIVHFGIRPAQ 781
  :::::::::::::::::::::
Db 684 MLOMEELEROKYAECKGMIVTSFGMKPSSVRIDQDREQFNDVITPFIIVHFGIRPAQ 743
QY 782 LSYAGDPQYQKLMKSYVKLRHLLANSPKYQDKOKLAOREALOKIROKNTMRREVTYE 841
  :::::::::::::::::::::
Db 744 LSYAGDPQYQKLMKSYVKLRHLLANSPKYQDKOKLAOREALOKIROKNTMRREVTYE 803
QY 842 LSSQGFVKTGIRSDVOQHAML.PVLTHNHRKHOCMLHDLKLGTYTODDCILQAMTHS 901
  :::::::::::::::::::::
Db 804 LSSQGFVKTGIRSDVOQHAML.PVLTHNHRKHOCMLHDLKLGTYTODDCILQAMTHS 863
QY 902 HHLNFGMNPDHANSLNSGIRQPKGDRKVMHMRKKGINTLIMISRLGDDPTPSR 961
  :::::::::::::::::::::
Db 864 HHLNFGMNPDHANSLNSGIRQPKGDRKVMHMRKKGINTLIMISRLGDDPTPSR 923
QY 962 INHNERLEFLGDAVVEFLTSVHL.YL.FPSLEEGGLATYRTAIVONOHMLAKLELDPF 1021
  :::::::::::::::::::::
Db 924 INHNERLEFLGDAVVEFLTSVHL.YL.FPSLEEGGLATYRTAIVONOHMLAKLELDPF 983
QY 1022 MLYAHBPDCREBDLHMAANCFEALIGAVYLEGSLBEAKOLFGRILLFNDPDLREVMY 1081
  :::::::::::::::::::::
Db 984 MLYAHBPDCREBDLHMAANCFEALIGAVYLEGSLBEAKOLFGRILLFNDPDLREVMY 1043
QY 1082 PLHPLLOBPNDROILETS.PVLQKLTPEFEALGVFTHVRLARAPLRTYVGFNLTIG 1141
  :::::::::::::::::::::
Db 1044 PLHPLLOBPNDROILETS.PVLQKLTPEFEALGVFTHVRLARAPLRTYVGFNLTIG 1103

```

```

QY 1142 HNORMEPLGDSIMOLVATEYLPIHRPDHHEGLTLRSLVNNRTOAKYAEELGMOEVAL 1201
  :::::::::::::::::::::
Db 1104 HNORMEPLGDSIMOLVATEYLPIHRPDHHEGLTLRSLVNNRTOAKYAEELGMOEVAL 1163
QY 1202 TNDKTKRPVGLRTKTLADLESFIALYTDKLEVYHFMVCFPPRLKEFTLNDMPNDP 1261
  :::::::::::::::::::::
Db 1164 TNDKTKRPVGLRTKTLADLESFIALYTDKLEVYHFMVCFPPRLKEFTLNDMPNDP 1223
QY 1262 KSQLQCCITLTTEGKEPDIPLYKTLQTVGPHANTYVAVYFKGERICCGGPSIQOAE 1321
  :::::::::::::::::::::
Db 1224 KSQLQCCITLTTEGKEPDIPLYKTLQTVGPHANTYVAVYFKGERICCGGPSIQOAE 1283
QY 1322 MGAAMDALKVNPOMAHOKRPIGRKROELKEMMREHOREPDETDIKK 1374
  :::::::::::::::::::::
Db 1284 MGAAMDALKVNPOMAHOKRPIGRKROELKEMMREHOREPDETDIEARK 1336

RESULT 5
Q80269_MOUSE PRELIMINARY; PRT; 1021 AA.
ID Q80269_MOUSE PRELIMINARY; PRT; 1021 AA.
AC Q80269;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Ribonuclease III.
GN Name=Etohl2; Synonym=Knaesen;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Kidney;
RX MEDLINE=22730743; PubMed=12191433;
RA Fortin K.R., Nicholson R.H., Nicholson A.W.;
RT "Mouse ribonuclease III. cDNA structure, expression analysis, and
RT chromosomal location."
RL BMC Genomics 3:26-26(2002).
DR EMBL; AF531013; AA065849.1; -; mRNA.
DR HSSP; O67082; IJF2.
DR Ensembl; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etohl2.
DR MGI; MGI:1261425; Knaesen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:004525; F:ribonuclease III activity; IEA.
DR GO; GO:006396; P:RNA processing; IEA.
DR InterPro; IPR001159; ds_RNA_bd.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS0142; RNASE_3_2; 2.
SQ SEQUENCE 1021 AA; 118200 MW; 15E4899270AF6F29 CRC64;

Query Match 69.6%; Score 5217; DB 2; Length 1021;
Best Local Similarity 73.1%; Pred. No. 1.5e-246;
Matches 1003; Conservative 10; Mismatches 8; Indels 352; Gaps 1;

```

```

QY 2 MGNNTCHMSFPRGRCGPRGSGHGAAPSFRPONTLLHPQDPVOYQYPPSAPST 61
  :::::::::::::::::::::
Db 1 MGNNTCHMSFPRGRCGPRGSGHGAAPSFRPONTLLHPQDPVOYQYPPSAPST 52
QY 62 TFSNBPAPNLPDRDPVFPMPMPPSAOGPLRPPCIRPPPNHQMRHPPVPPCPPMP 121
  :::::::::::::::::::::
Db 53 ----- 52
QY 122 PMPPCPNNPVPGAPPGQGTFFPMMPRPMRPPPPVPMQVUNYQVPPGYSHHNPSPS 181
  :::::::::::::::::::::

```

```

Db 53 ----- 52
Qy 182 FNSFONNPFSLPSANNSSPHFRHLPPYPLPKAPSEERSPERLKAYDDHRRHSHGRG 241
Db 53 ----- 52
Qy 242 EHRSLDRRGRSPDRRRQDSRYSDYRGRTPSHRSYENSREPERHRHRDRNSP 301
Db 53 ----- 52
Qy 302 SLERSYKKEYSRGREYGLVPEPACCTPELPGELIKNTDSWAPPLELVNHRSPREKK 361
Db 53 ----- 52
Qy 362 RARWEEKDRWSDNQSGDKNYTSIKEKEPEETMPDKNEEBEELKPMWIRCTSENY 421
Db 53 -----ELKKVMIRCHSEXY 68
Qy 422 YSSDPMDOVGSDTVGTSLRLDLYDKFEEELSGROKAKAAPPMWPPKTKLDEDESSS 481
Db 69 YSSDPMDOVGSDTVGTSLRLDLYDKFEEELGNROKAKAAPPMWPPKTKLDEDESSS 128
Qy 482 EEECESDEBTSSSSSDSSEVPVIAIKKKKAPDLHDLWYNDPGQMNDEPLCKCSAK 541
Db 129 ESECETDSDTSSSSSDSSEVPVIAIKKKKAPDLHDLWYNDPGQMNDEPLCKCSAK 188
Qy 542 ARTGTGRHSYGEBAIKPCRPMTNNAAGRLFHYRIIVSPPTNPLDRPTVIEYDHEXYF 601
Db 189 ARTGTGRHSYGEBAIKPCRPMTNNAAGRLFHYRIIVSPPTNPLDRPTVIEYDHEXYF 248
Qy 602 EGSFMAAPLNTNPLCKVIRNIDYTIHFIEBEMPEFCVKGLELSLFLFRDILEYD 661
Db 249 EGSFMAAPLNTNPLCKVIRNIDYTIHFIEBEMPEFCVKGLELSLFLFRDILEYD 308
Qy 662 NWLKGPLFEDSPCCRFHMPRFVFLFDGGEVLSMHOILLYLRCKSALVPEBEIAN 721
Db 309 NWLKGPLFEDSPCCRFHMPRFVFLFDGGEVLSMHOILLYLRCKSALVPEBEIAN 368
Qy 722 MLOWEELEWOKYAEBCKGMIVTNPGTSPSSVRIDQDRQOFNEVTFPIIVHFGIRPAQ 781
Db 369 MLOWEELEWOKYAEBCKGMIVTNPGTSPSSVRIDQDRQOFNEVTFPIIVHFGIRPAQ 428
Qy 782 LSYAGDPOYOKLMSYVYKLRHLANSPKVYQTDOKOQAOBEALOKIROKNTMRREYVE 841
Db 429 LSYAGDPOYOKLMSYVYKLRHLANSPKVYQTDOKOQAOBEALOKIROKNTMRREYVE 488
Qy 842 LSSQGFWKGTGIRSDVCQAHMMLPVLTHTIRYHOCMLHDLKLGTYFQDRCLLOLANTHPS 901
Db 489 LSSQGFWKGTGIRSDVCQAHMMLPVLTHTIRYHOCMLHDLKLGTYFQDRCLLOLANTHPS 548
Qy 902 HHLNFGMNDHARNSLNSGIRQPKYGDRCVHHMHMRKKGINTLINIMSRLODDPTSPR 961
Db 549 HHLNFGMNDHARNSLNSGIRQPKYGDRCVHHMHMRKKGINTLINIVSRLODDPTSPR 608
Qy 962 INHNRELEFGDAVVEFLTSVHLYYLPSELESGGLATYRTAIVONQHLMLAKKLELDF 1021
Db 609 INHNRELEFGDAVVEFLTSVHLYYLPSELESGGLATYRTAIVONQHLMLAKKLELDF 668
Qy 1022 MLYAHGPDICRESDDLHMANCFEALIGAVYLEGSLEBAKQFGRLLFPNDPGLREWMLNY 1081
Db 669 MLYAHGPDICRESDDLHMANCFEALIGAVYLEGSLEBAKQFGRLLFPNDPGLREWMLNY 728
Qy 1082 PLHNPLOQEPNTRDOLIEFSPVLOKTEFEBAIGVFTHVRLLAARFTLRTVGPHNLTLTG 1141
Db 729 PLHNPLOQEPNTRDOLIEFSPVLOKTEFEBAIGVFTHVRLLAARFTLRTVGPHNLTLTG 788
Qy 1142 HHQRMSEFLDSDIMQVATERYLFIHFPPDHGHLTLRSGLVNNRTOAKAAEELGMOEYAV 1201
Db 789 HHQRMSEFLDSDIMQVATERYLFIHFPPDHGHLTLRSGLVNNRTOAKAAEELGMOEYAV 848
Qy 1202 TNDKTKRPVGLATKTLADLLESFIALYTDKOLEVYVHTFMVNCFFPRKPEFILNDMPND 1261
Db 849 TNDKTKRPVGLATKTLADLLESFIALYTDKOLEVYVHTFMVNCFFPRKPEFILNDMPND 908

```

```

Qy 1262 KSQLOQCCLTTRTEGKEKPDIPLYKTLQTVGSPSHARTTYAVYFKGERICGKGPSIQOAE 1321
Db 909 KSQLOQCCLTTRTEGKEKPDIPLYKTLQTVGSPSHARTTYAVYFKGERICGKGPSIQOAE 968
Qy 1322 MGAANDALEKYNPQMAHQKRTIGKRYROELKEMERERERHORPEPETDIDK 1374
Db 969 MGAANDALEKYNPQMAHQKRTIGKRYROELKEMERERERHORPEPETDIDK 1021

RESULT 6
Q6PF8_MOUSE
ID Q6PF8_MOUSE PRELIMINARY; PRT; 942 AA.
AC Q6PF8;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Naesen protein (Fragment).
GN Name=Etohl2; Synonyms=Knaesen;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Czech II;
RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.24260899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner I., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield V.S.N., Krzywinski M.J., Skalska U., Smailus D.B.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Czech II;
RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RG NIH MGC Project;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC057687; AAH57687.1; -; mRNA.
DR Ensembl; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etohl2.
DR MGI; MGI:1261425; Knaesen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; P:double-stranded RNA binding; IEA.
DR GO; GO:0004525; P:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; D9_RNA_bd.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00366; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS50142; RNASE_3_2; 2.
FT NON_TER 1 1

```



```
QY 715 PEEBIAANMLQWEELEMOXYAECKGMIVTNPGRKPSVVIDOLDREOFNPDVITPPIIYN 774
|||
DB 121 PEEBIAANMLQWEELEMOXYAECKGMIVTNPGRKPSVVIDOLDREOFNPDVITPPIIYN 180
QY 775 FGRPAQOLSVAGDPQOYOKLWKSYYVYKLRHLANSRKYQDQKQDLAOREEALQIKIRQKNTM 834
|||
DB 181 FGRPAQOLSVAGDPQOYOKLWKSYYVYKLRHLANSRKYQDQKQDLAOREEALQIKIRQKNTM 240
QY 835 RREVTVELSSQGWKYGIRSDVQCHAMMLPVLTHTIRYQCSLHNDKLGTYTQDSCDQ 894
|||
DB 241 RREVTVELSSQGWKYGIRSDVQCHAMMLPVLTHTIRYQCSLHNDKLGTYTQDSCDQ 300
QY 895 LAMTHPSHHLNFGMNDHANSLNSCGIRQPKYDKRVHNMNRKKGINTLINIMSRLQ 954
|||
DB 301 LAMTHPSHHLNFGMNDHANSLNSCGIRQPKYDKRVHNMNRKKGINTLINIMSRLQ 360
QY 955 DDPTRIRINHNRELEFLGDAVVEFLTSVHLUYLPESLBSGLATTYTAIVONHMLAK 1014
|||
DB 361 DDPTRIRINHNRELEFLGDAVVEFLTSVHLUYLPESLBSGLATTYTAIVONHMLAK 420
QY 1015 KLELDPFMLYAHGPDLCRESDDLHMANCPREALIGAVLEGSLEBAKOLFGRLLFNDPDL 1074
|||
DB 421 KLELDPFMLYAHGPDLCRESDDLHMANCPREALIGAVLEGSLEBAKOLFGRLLFNDPDL 480
QY 1075 REVWMLNPLRPLQLOBPNTDROLIETSPVLOKLTREBEALGVIFTHVRLARAFILRTVG 1134
|||
DB 481 REVWMLNPLRPLQLOBPNTDROLIETSPVLOKLTREBEALGVIFTHVRLARAFILRTVG 540
QY 1135 FNNHLLTGNHQRMEFLDSTIMOLVATYTLPHRPDHHGHLLTLRSSLVNNRTOAKYAEEL 1194
|||
DB 541 FNNHLLTGNHQRMEFLDSTIMOLVATYTLPHRPDHHGHLLTLRSSLVNNRTOAKYAEEL 600
QY 1195 GMOEYAITNDKTRRPGVGLRTKTLADLESFIALYDKOLEVYHTPMNCFPRLLKEFTL 1254
|||
DB 601 GMOEYAITNDKTRRPGVGLRTKTLADLESFIALYDKOLEVYHTPMNCFPRLLKEFTL 660
QY 1255 NODMNDPKSOLOOCCLTLRTEGKEPDIPLYKTLQTVGSPSHARTYVAVYFKESRIGCGKG 1314
|||
DB 661 NODMNDPKSOLOOCCLTLRTEGKEPDIPLYKTLQTVGSPSHARTYVAVYFKESRIGCGKG 720
QY 1315 PSIOQAEMGAMDALEKYNPQMAHQKRFIERKYRQELKEMERERHQBEPEDEDIK 1374
|||
DB 721 PSIOQAEMGAMDALEKYNPQMAHQKRFIERKYRQELKEMERERHQBEPEDEDIK 780

RESULT 8
Q4SKM1_TETNG PRELIMINARY; PRT; 1041 AA.
AC Q4SKM1;
DT 13-SEP-2005 (Tremblrel. 31, Laet sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Laet annotation update)
DE Chromosome undetermined SCAFI4565, whole genome shotgun sequence.
ORFNames=GSTENG0016530001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
[1]
RN NUCLEOTIDE SEQUENCE.
RP Jallou O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Flaher S., Lutfalla G., Dossat C., Segutens B.,
RA Desliva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Casseletti V., Katinka M., Vachele B.,
RA Blemont C., Skellin Z., Catolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lander G., Chapuis C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Westrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
```

```
RA Laudet V., Schachter V., Quetier F., Saurin M., Scarpetti C.,
RA Winkler P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ database.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEB014565; CAF98721.1; -; Genomic DNA.
SQ SEQUENCE 1041 AA; 120534 MW; 2CA7A165A7BB3CDB CRC64;

Query Match 50.2%; Score 3767; DB 2; Length 1041;
Best Local Similarity 67.5%; Pred. No. 1.1e-175;
Matches 728; Conservative 80; Mismatches 135; Indels 140; Gaps 18;

QY 10 MSFHPGCGPRGCGHAGAPSPAPSPQNLRLHPQOPVOXYGEPSPASTTF-SNSPA 68
|||
DB 1 MSFHPDRAHAPGMAAPGPMSPQNYRPTLRRPVPVRYHYDQVPPSSGYHNS-- 58
QY 69 PNFLLPRDPFVPPPP-----PMPSAQP-LPPCPTLRPPRNQMRHFFVPPPC 116
|||
DB 59 -GYMHPSTFMQYSPGAPADSRSEPPVAPQEGPRLPCLPRTPTSLYMMHPVPPPLA 117
|||
QY 117 PPMPPMPCPN---NPPVGAAPGCGTFPPMMRPSMHPPPPPVPPQVNVYQPPGYS 173
|||
DB 118 PPMPPMPSPSSTHPPY-----PMSHPHPP----- 147
QY 174 HNNFPPSPFNS-FQNNPSSFLPSANSSPHRNLPPVPLPKAPSERSPERLKHVDHR 232
|||
DB 148 ---FPPPTNSGHHVQSGSFAFDLG-----FRHGSPYHGXDKPADKRRRPPGYRHDLR 198
QY 233 HDHGH-GEGERHSL---DRBERGSPDR--RQDSKRSYDGRGTRSRHSYKRSR 286
|||
DB 199 OGYSHSGHSDQKMFPPGDRKDRGRSPRRMRPEGGRHSEYDRERTPRH---RSRE 254
QY 287 RREBRHHRDNRSPSLERSYKKEYRSGRSGSLVPPRAGCTPLPGBIINKDMSWR 346
|||
DB 255 RREBRHHRDNRSPSLERSYKKEYRSGRSGSLVPPRAGCTPLPGBIINKDMSWR 283
QY 347 PLEIVNHRSPSEKRAARWEEKDMSDQSS-GKDNYTSIKEPEETMPDK----- 399
|||
DB 284 -----SRERKSRWEERERSESSAPSRKRSYVSARNRSEBAVPEREGCPR 333
|||
QY 400 -----NEEBEELLKPVWIRCTHSENYSSDPMDOYGDSTVGT 438
|||
DB 334 TRGREGEPALPEHDEKEKEGEKEEELLKAWIRCTHAESYSSNDPMDOYGDSTVGT 393
QY 439 SLRLDYDKFEELSGROKAKAARPPWPPPTKIDEDL-ESSSEBECSDSD-STCSS 496
|||
DB 394 SKLRDLYORFEELGRRQARAARPPPTKIDEDLDESSSEBECSDSDAGSTCSS 453
QY 497 SDSEVDFVIAELKRAKAPDRILHDELAMVNDPQOMDGPCKCSAKKARFRGHSIYPGE 556
|||
DB 454 SDSDVDFVIAELKRAKAPDRILHDELAMVNDPQOMDGPCKCSAKKARFRGHSIYPGE 513
QY 557 AIKPCRPMTNNAQRLFNHYRITVSPPTNFLTDRPVIIEYDHEVIFEGFSMAHAPLTNP 616
|||
DB 514 SVKQCRAMNNNAQKLFHYRITVSPPTNFLTDRPVIIEYDHEVIFEGFSLTPLTNP 573
QY 617 LCKVIRENDYTHIFEEBMPENFCYKGLFSLPFRITLTYMNLGRLPEEDSPCC 676
|||
DB 574 LCRVIRENDYTHIFEEBMPENFCYKGLFSLPFRITLTYMNLGRLPEEDSPCC 633
QY 677 PRFHMPRFRVRLPDGKGVLSMHOQLYLRLCSKALVBEERIANMLQWEELEMOXYAE 736
|||
DB 634 PRFHMPRFRVRLPDGKGVLSMHOQLYLRLCSKALVBEERIANMLQWEELEMOXYAE 693
QY 737 CGKMITVNGTPSSVVIDOLDREOFNPDVITPPIIYHFGIRPAQOLSVAGDPQOYOKLWKS 796
|||
```

Db 694 CKGMIVTNGMKPSSVRIDQLDRBOFNPVITFPITIVHFGIRPAQLSYAGDPQYOKLWKS 753
QY 797 YKCLRHLLANSPKVKQTDKOKLAQREBALOKIRKQNTMRREVTVELSSGFWKGTGRSDY 856
Db 754 YKCLRHLLANSPKVKQTDKOKLAQREBALOKIRKQNTMRREVTVELSSGFWKGTGRSDY 813
QY 857 COHAMMLPVLTHHIRHYOCLMHLIDKLGTYTQDRCLLQOLAMTHPSHHILFGNMPHANS 916
Db 814 COHAMMLPVLTHHIRHYOCLMHLIDKLGTYTQDRCLLQOLAMTHPSHHILFGNMPHANS 873
QY 917 LSNCGIRKQYDKRKVHHNMRKKGINTLINIMSLRGODDPSRINNHRELEPLGDAYV 976
Db 874 LSNCGIRKQYDKRKVHHNMRKKGINTLINIMSLRGODDPSRINNHRELEPLGDAYV 933
QY 977 EFLTSVHLVYLPFSLBEGGLATYRTAIYONOHMLAKKLELDPEMLYAHGPDLCRESDL 1036
Db 934 EFLTSVHLVYLPFSLBEGGLATYRTAIYONOHMLAKKLELDPEMLYAHGPDLCRESDL 993
QY 1037 RHAMANCFEALIG 1049
Db 994 RHAMANCFEALIG 1006
RESULT 9
Q80U5 MOUSE PRELIMINARY; PRT; 705 AA.
ID Q80U5 MOUSE PRELIMINARY; PRT; 705 AA.
AC Q80U5-
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Name=Ecch12; Synonym=Rnaseen;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
[1]
NM NUCLEOTIDE SEQUENCE.
RP STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
Dietzenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.U., Ustin T.B., Tobin L.S., Carninci P., Frange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Whiting J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Blakesley R.C., Tuchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalms D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
NM NUCLEOTIDE SEQUENCE.
RP STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;
RC NIH MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC050057; AAH50057.1; -; mRNA.
DR HSSP; 067082; 1JFZ.
DR Ensemble; ENSMUSG0000022191; Mus musculus.
DR MGI; MGI:1261425; Etch12.
DR MGI; MGI:1261425; Rnaseen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.

DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; Ds_RNA_bd.
DR InterPro; IPR000999; Rnase_III.
DR Pfam; PF00035; dsm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00358; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; Rnase_3_1; 2.
DR PROSITE; PS50142; Rnase_3_2; 2.
FT NON_TER
SQ SEQUENCE 705 AA; 82084 MW; 332C69A5F1977D6 CRC64;
Query Match 49.3%; Score 3696; DB 2; Length 705;
Best Local Similarity 98.7%; Pred. No. 2e-172;
Matches 696; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 670 EDSPPCCRRFHFMPFVFLPDGKGVLSMGIILYLRCGKALVPEEIANMLQMELE 729
Db 1 EDSPPCCRRFHFMPFVFLPDGKGVLSMGIILYLRCGKALVPEEIANMLQMELE 60
QY 730 WOKYAECKGMIVTNGMKPSSVRIDQLDRBOFNPVITFPITIVHFGIRPAQLSYAGDPQ 789
Db 61 WOKYAECKGMIVTNGMKPSSVRIDQLDRBOFNPVITFPITIVHFGIRPAQLSYAGDPQ 120
QY 790 YOKLWKS YKCLRHLLANSPKVKQTDKOKLAQREBALOKIRKQNTMRREVTVELSSGFWK 849
Db 121 YOKLWKS YKCLRHLLANSPKVKQTDKOKLAQREBALOKIRKQNTMRREVTVELSSGFWK 180
QY 850 TGISRDVCOHAMMLPVLTHHIRHYOCLMHLIDKLGTYTQDRCLLQOLAMTHPSHHILFGN 909
Db 181 TGISRDVCOHAMMLPVLTHHIRHYOCLMHLIDKLGTYTQDRCLLQOLAMTHPSHHILFGN 240
QY 910 PDHANSLSNCGIRKQYDKRKVHHNMRKKGINTLINIMSLRGODDPSRINNHRELE 969
Db 241 PDHANSLSNCGIRKQYDKRKVHHNMRKKGINTLINIMSLRGODDPSRINNHRELE 300
QY 970 FLGDAVVEFLTSVHLVYLPFSLBEGGLATYRTAIYONOHMLAKKLELDPEMLYAHGPD 1029
Db 301 FLGDAVVEFLTSVHLVYLPFSLBEGGLATYRTAIYONOHMLAKKLELDPEMLYAHGPD 360
QY 1030 LCRESDDLPHANANCFEALIGAVYLEGSLBEAKQLFGRLLFNDPDLREWLANPLPLQIQ 1089
Db 361 LCRESDDLPHANANCFEALIGAVYLEGSLBEAKQLFGRLLFNDPDLREWLANPLPLQIQ 420
QY 1090 EBNTRQRLIETSPVLQKLTPEFEALIGVFTYHRLARAFLLRTVGFNNHTLGHNMREFL 1149
Db 421 EBNTRQRLIETSPVLQKLTPEFEALIGVFTYHRLARAFLLRTVGFNNHTLGHNMREFL 480
QY 1150 GDSINQVATEVFLFHPDHEGHLTLRSSLVNNRTQAKVAELGMOEYATINDTKRP 1209
Db 481 GDSINQVATEVFLFHPDHEGHLTLRSSLVNNRTQAKVAELGMOEYATINDTKRP 540
QY 1210 VGLRKTTLADLLESTIALYTDKLELYHTFNANVCFPRPKFPIINQDNDKSQLQOCC 1269
Db 541 VALRKTTLADLLESTIALYTDKLELYHTFNANVCFPRPKFPIINQDNDKSQLQOCC 600
QY 1270 LTLRTEGKEPDIPLKTTQTVGSHARTYTVAVYKGBEIGGCKPSIQAEAGAMDAL 1329
Db 601 LTLRTEGKEPDIPLKTTQTVGSHARTYTVAVYKGBEIGGCKPSIQAEAGAMDAL 660
QY 1330 EKYNFPQMAHQKRFYGRKYRQELKEMWEREHOEREPDETEDIK 1374
Db 661 EKYNFPQMAHQKRFYGRKYRQELKEMWEREHOEREPDETEDIK 705
RESULT 10
Q8XN5 DROME PRELIMINARY; PRT; 1327 AA.
ID Q8XN5 DROME PRELIMINARY; PRT; 1327 AA.
AC Q8XN5-
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)


```

Db 47 SSKSLDYVYPETPAYASS---VPSYDPYQPPAYGEGYAVNEQAQKYGQESHYQYQ 103
QY 180 PSFNFONNPSFELPSANNSSSPHRLPPYPL----PKAPSERRSPELTKHYDDHRD 235
Db 104 PA-----SGSFLYES-----XYPRYPAYSNNRPSERO-----RTYS 139
QY 236 HSHGGERHRSIDRRERGRSPDRRQDSRYRSDYDGRGTPSHRSYSERERERERHRR 295
Db 140 NSSSGYHHYF-----GYSSGR-----RYEQRHQO-----EHRQIDSRVYHPRHGAY 183
QY 266 DNRBRPSLSRYKKEYKSGSGSYGSLVVERPAGCTPELBGELIKNTDSPAPELTVNRS 355
Db 184 AHRQAGSQGHY---YGSAAFN-----YQSDVDSYSGHHEREKN 219
QY 356 PEREKRA--RMEBEKDR-----MSDNQSGDKNYTSIKEKEPEETMPDKNEEBELIK 409
Db 220 ETLKTRAPKQYETERDLRLRQWCSNFC-----EKPEDEVKGMALSEDAV 267
QY 410 PWMIRCTHSENYSSDPMDQVDSVTVGTSTRLDLYDKFEELSGROEKAKARPEWEP 469
Db 268 BSWAVSSPAPELYERTK-----SENEVRGARLOKCTLPDEELLQRAKVRKELPYVVP 323
QY 470 KTYL-----DBDESSSESCESEDSTCSSSDSEVPDV-----IAEIKRKAN 514
Db 324 PKKARRRYCKKHKKSEACSSSSSSDDSDDEDA-----FKLEQDCMBELSKYQH 373
QY 515 PRLDELAYNPDGONNDGPLCKCSAKARTGIRHSIYPGEBAIKPCRPMTNAGRLPAY 574
Db 374 PQRVHADLWMDAGENNDGPLCKCSAKSRIGIRHGIYGEYGYKLCDSNNAGLFLFY 433
QY 575 RTVSPTNFLTDRPTVIEYDDEHYIEGFSMFAAPLTNIPLCYIRFNIDYTHFIEE 634
Db 434 RISISPTNFLTKTPTIIEKHDEHEFLFEGFSLSHVRLSDLPVCYIRFNIEYTHEBE 493
QY 635 MPENPCVKGLEFSLFLRDLIELYDNLKGPL----FEDSPCCRRHFMFRPRFLP 690
Db 494 KMPENFTIHELDIIFFKYLFHELELDVFNLMPLPSGNVEES---CPAHFFPRFRDLP 550
QY 691 DQGEKYLASHOILLYLRCSKALVPEEELANMLQMELEEMOKYAECKGMITNPGTKS 750
Db 551 DNGKEVLAVVEYRILYLDNSAQULVEQOILHLANOISQEMQYVDITKMLYTKPKYKC 610
QY 751 SVRIDQLDREOFNPDV-----ITPPLVHFGIRPAQLSYAGDPQYQKLMKSYVKLR 801
Db 611 SLRVQDLDN--NSDLPECVDRGTGSHRALVHFGIHPQLSYAGPEYQKAMREYVKYR 668
QY 802 HLLANSFKYQDQKQKLAOREELQKIRQKQNTREYVYELSSQGWKIGISDVQHAM 861
Db 669 HLMANMSKSPFKDKRKLBEKEQRLQEMRTOGRMKRNIYVAISSEGFYRTGIMCDVVQHAM 728
QY 862 ML.PV.LTHHTRHOCMLHDKLIGYTFODCLQLQAMTHPSHHLNFCMNPDPHANSLSNG 921
Db 729 LIPVLGHLRFHNSLDLLESISYIRKRYLQDLALITHSYKKNYGTNDPHARNSLTNG 788
QY 922 IQPKYGDGRVHMHMRKKGINTLINIMSRIGODDTPSRINNELEFLGDAVVEFLTS 981
Db 789 IQPEBGDKRIHMYNRKGINLVGIMSRFGKEHEFVSNTTNERLEFLGDAVVEFLSS 848
QY 982 VHLVYLPESLBEGGLATYRTAIYQNOHMLAKKLELDPMLYAHNPDLCSREDLRHMA 1041
Db 849 IHLFFMFPELEBEGGLATYRAAIYQNOHMLAKKLELDPMLYAHNSDLCHELELHMA 908
QY 1042 NCEALIGAVYLEGSLEBAKOLFGRLLF--NDPDLREVMYLPYLPLOLEPNTDQLIEF 1100
Db 909 NCEBALMGALLDGGIKVADVEFTDALFRODEKLISIKNLPHPHLOBOEPDLRSCIDS 968
QY 1101 SPYLQKLFEEBAIGYIFTHVRLIARAFLLTVGFNHLTLGNORMEFLDSTIMQVATE 1160
Db 969 YARLKLTLKFEDSIGIKFKHILARAFDRSIGFHLTLGSGNQLBFLGDVTLQILCSB 1028
QY 1161 YLFIHPDDHHEGHLTLRSLVNNRTOAKYAEELGMOEYAI--TNDTKRPAVGARTFLAD 1219
Db 1029 YLRRHPRHEHGHSLIRSSLVNNRTOAVVCDLGMPCYAVVANPK---ADLKTDRAD 1084

```

```

QY 1220 LLESFIALYTDKDLVYHTMNVNCFPPRLKEFLINDMNDPKSLOOCCLTLLRT--EKE 1278
Db 1085 LLEAFGLAYVDKGLLYCEOFCHVCLFPLQLFLINNDQWNPBKSLOOCCLTLLRTMDGE 1144
QY 1279 PDILPYKTLQVGPESHATYVAVYFKGERIGCGGPSIQOAEWGAADALEKYN--FPQ 1336
Db 1145 PDIPYKVAEASGPPNTVYKAVYFRSKRLATSSGSSIIQOAEWGAALENSRLDFPO 1204
QY 1337 MAHQRFPIGRKYRQEL---KEWRWERHQBEPDETDIKK 1374
Db 1205 LDHQKRVIAKSKIKQGTGNELNDSDRQHJ-----EKIKR 1239

```

RESULT 11

OSTOK8 ANOGA PRELIMINARY; PRT; 979 AA.

AC 05TOK8, 01-FEB-2005 (Tremblrel, 29, Created)
 DT 01-FEB-2005 (Tremblrel, 29, Last sequence update)
 DT 01-FEB-2005 (Tremblrel, 29, Last annotation update)
 DE ENSANGP0000028563 (Fragment).
 GN ORFNames=ENSANGG0000009133;

OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 ON Anopheles; Anopheles.
 NX NCBI_Taxid=180454;

NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RT "Anopheles gambiae re-annotation."
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

CC EMBL; ABA01008964; BA139656.1; Genomic_DNA.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0004525; F:ribonuclease III activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR InterPro; IPR001159; DS_RBD.
 DR InterPro; IPR000999; RNase_III.
 DR Pfam; PF00035; dsm; 1.
 DR Pfam; PF00636; Ribonuclease_3; 2.
 DR SMART; SM00358; DSRM; 1.
 DR SMART; SM00355; RIBOC; 2.
 DR PROSITE; PS00137; DS_RBD; 1.
 DR PROSITE; PS00517; RNASE_3_1; 1.
 DR PROSITE; PS0142; RNASE_3_2; 2.
 FT NON TER 1
 FT NON TER 979
 SQ SEQUENCE 979 AA; 113068 MW; 46F4AE2A0257AD92 CRC64;

Query Match 38.4%; Score 2861.5; DB 2; Length 979;
 Best Local Similarity 55.3%; Pred. No. 1,8e-132;
 Matches 546; Conservative 169; Mismatches 235; Indels 37; Gaps 12;

```

QY 382 KNYTSIKKEPPEITPMDKNEEBELIKVWIRCTHSENYSSDPMDQVGDSTVVGNSRL 441
Db 13 RNYCETSEDIAKLKEMANDED-----RTLWVSSPAELIYK-----RVS DKVVESTAKL 62
QY 442 RLLYDKFEELSGROEKAAARPPWPPKTKLDEDESSSESCESEDSTCSSSDSEY 501
Db 63 DALCTLFEBELIKRAERITATGTPTVNPBPKKMKCKRHKHXC-----SSSESSEDEEM 117

```


Qy	502	FD--VLAETKRKAANDRLHDELAMVNDPQGMNDGPLCKGAKARRGIRGIRHSIYRGEA	557
Db	118	EFEDSCSMEELTAKIKHPIRLHVLMDVMDPGMDNGPLCRCSARRRRTGIRGKXPGSEG	177
Qy	558	IKPCSPMNNAAGRLPHYRITVSPRPNNFLDRFTVIEYUDHEXIFPGFSWFAAPLTINPL	617
Db	178	FPKCI.PNSNNADKLHYHRTITSPNPFLTKTTIHKHOHELEFEGFSILAHPIGELPT	237
Qy	618	CKVIRFNI.DYTIHFIEEMMPENFCVKGDELFSLPFRDILBELVDNMLKGPLFEDSPCCP	677
Db	238	CKVIRFNI.EYTLIVIEEQMPENFTRI.RELNLPRYIFRELBELVDPTVQ--PSSGSDSCP	296
Qy	678	RPHFPRPRVRLPDGCKEYLSMHQILLYLRSCKALVBEELANMLQWELDEWQYAEBC	737
Db	297	CYHFLPRFVRDL.PDNKGEVLAHMEVTLARYL.DNSGFLVPPDMLKEMMDMSQNMEDYVUV	356
Qy	738	KGMITNNGTSPSSRVIDOLD-----EGFNDDVTPPIIVHFGIRPOLSTVAGDQ	789
Db	357	KGMVSNNGMKPCSRVOLDNNDVDBANAI.DSGLVHPRTVHFGIRPOLSTVAGNPE	416
Qy	790	YQKLKSYVULRHLLNLANPKVQOTDQKQLAQREBALQKIROKNTARRREVTELSQCGPWK	849
Db	417	YQKAREVIEKRRHLLANNSKSPFEDKRLKAEKENLLEMRQNGRKNTITAVSKAKHR	476
Qy	850	TGIRSDVCQAHMMLVLTNHIRYHOCLMHLDKLIGYTFQDRCLLOLANTHPSHHLNFGMN	909
Db	477	TGIMCDMWQAHMLLPVLTGHLRFHRSLSVLVEKYIETFTNRYTLOLALTHPSYKENPSTN	536
Qy	910	PDHARNSLSNGCIRPKXGDRKVHNMHRKKGINFLINMSRLGDDPTPRIRINNEBLE	969
Db	537	PDHARNSL.NGCIROPREYGDRIKHMTNRRKGINFLISIMSPFGKEHETDSITINERLE	596
Qy	970	FLGDAVVEFLT.SVHLUYU.FPRLSEEGSLATYRATIAQONH.LAMAKLELDPFMLYAHQPD	1028
Db	597	FLGDAVVEFLT.SIHLFHM.FPRLSEEGSLATYRATIAQONH.LAMAKLELDEPMLYAHSD	656
Qy	1030	LCRESBLRHMANCFEALIGAVLYEGSLSEAKOLFGLLFNDPD--LREWVNIYPLRLQL	1088
Db	657	LCHELELRLHALANCFEALMGALLDGGIEVADRVAYAL.FQEDDTLRGIWVYPSHPLOE	716
Qy	1089	QEPNTDROLIETSPVLOKLTREBEALGYITFTVRLAAAFPLRTYGFPHMLTIGHNQMEF	1148
Db	717	QEPNRDRHHIS.FEMLKTLTRFEDSIGVQFNHILRLAAFPDRSIGFNLTLGSGNORLEF	776
Qy	1149	LGD.SIMOLVATELYE.IHFRDHEGHLTLRSSLVNRRQAKVABEELGMOEVAL--TNDCTK	1207
Db	777	LGDVLYLOLIGSELYRHF.PEHNEGHLSLRSLSLVNRRQAVACDLGMYOAVVSNPK--	834
Qy	1208	RPVGLRTKTL.LDLESFPAALYTDLDLEYVTFEMVCFEPPLKEFILT.LQDNMDPKSOLOQ	1267
Db	835	--ADLKTDRADLLEAF.LGALIVYDKGLEYSCEMFCHVCLF.PPLQDRIMNQDNMDPKSKLOQ	892
Qy	1268	CCLTILRT--EGKEPDIPLKYTLQTVGPSHARTYTAAYVFKGIRIGCGKPSIQOEMGAM	1326
Db	893	CCLTILRTDGGEPDIPVYKVIETGEPNTIRVYSAVVYFRGRLACADGHSIQOEMMAAK	952
Qy	1327	DALF--KNFPOMAHQKRFGRKXQOE	1351
Db	953	QALENSKDLFPOLDHQRRVIAQSLKRO	979
RESULT 12			
Q8CJ74_MOUSE PRELIMINARY; PRT; 541 AA.			
AC Q8CJ74;			
DT 01-MAR-2003 (Tremblrel. 23, Created)			
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)			
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE Ribonuclease III			
GN .Name=Ech12; Synonym=Kns3, Naa5en;			
MS Mus musculus (Mouse).			
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			

OC Muroidea, Muridae, Murinae, Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Min K.-L., Galarnau A., Parniak M.A.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF440766; AAN63637.1; -, mRNA.
DR HSSP; O67082; 1HPZ.
DR EnsEmbl; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etoh12.
DR MGI; MGI:1261425; Nasen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; Ds RNA bd.
DR InterPro; IPR00999; RNase_III.
DR Pfam; PF00035; dsm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS0137; Ds RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS0142; RNASE_3_2; 2.
SO SEQUENCE 541 AA; 63007 MW; DQACG52A0CE044F1 CRC64;

Query Match	37.7%	Score 2825	DB 2	Length 541
Best Local Similarity	90.3%	Pred. No. 5e-130		
Matches 532	Conservative 3	Mismatches 6	Indels 0	Gaps 0
QY	MRREYVEISLSQGWKTKGIRSDVQCHMMMLVLTGHIIRYHOCALMLDCLIGYTERQDRCLL			893
DB	1 MRREYVEISLSQGWKTKGIRSDVQCHMMMLVLTGHIIRYHOCALMLDCLIGYTERQDRCLL			60
QY	QLAMTHPSHLNFGMNPDPHANSLSNCGIRPKYGDKRKVHMHMRKKGINTLINIMSRIG			953
DB	61 QLAMTHPSHLNFGMNPDPHANSLSNCGIRPKYGDKRKVHMHMRKKGINTLINIMSRIG			120
QY	954 QDDTPSRINNNELEFLGDVVFELTSSVHLXYLPSPSLEEGGLATYTPALYQNOHMLA			10133
DB	121 QDDTPSRINNNELEFLGDVVFELTSSVHLXYLPSPSLEEGGLATYTPAMVQNOHMLA			180
QY	1014 KKLGLDPRMLYAHGPRDLCRESDLRHMANCFEALIGAIVLEGSLEBAKQLRGLLENDPD			10737
DB	181 KKLGLDPRMLYAHGPRDLCRESDLRHMANCFEALIGAIVLEGSLEBAKQLRGLLENDPD			240
QY	1074 LREVLNLYPLHPLDLOEPNTDROLIETSPVLOKLTPEEALIGVIFTHVRLHARAFLLTV			11333
DB	241 LREVLNLYPLHPLDLOEPNTDROLIETSPVLOKLTPEEALIGVIFTHVRLHARAFLLTV			300
QY	1134 GFNLTLGHNORMFEGDSIMQVATVEYLFTHPPDHHGHGTLIRSSLVNNRTQAKVAEE			11939
DB	301 GFNLTLGHNORMFEGDSIMQVATVEYLFTHPPDHHGHGTLIRSSLVNNRTQAKVAEE			360
QY	1194 LGMOEVALITNDKTRPVGLRTKTLADLLESTIALYTDKOLEVHTMNVCFFPRLIKEFI			12533
DB	361 LGMOEVALITNDKTRPVGLRTKTLADLLESTIALYTDKOLEVHTMNVCFFPRLIKEFI			420
QY	1254 LNMOMNPQKSLQOCCCLTLRTGEGEPDIPVKTQLQTGSPSHARTYTVAAVYFKGIRIGCGK			13133
DB	421 LNMOMNPQKSLQOCCCLTLRTGEGEPDIPVKTQLQTGSPSHARTYTVAAVYFKGIRIGCGK			480
QY	1314 GPSIQOAEKMGAMDALEKTNFPOMAHOKRPIGRKTYROELKEMRERBHOERPEDETIDK			13733
DB	481 GPSIQOAEKMGAMDALEKTNFPOMAHOKRPIGRKTYROELKEMRERBHOERPEDETIDR			540
QY	1374 K 1374			
DB	541 K 541			

```
Q960Y4 DROME
ID Q960Y4 DROME PRELIMINARY; PRT; 1071 AA.
AC Q960Y4
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE LD29995p.
GN Name=drosha; ORFNames=CG8730;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carleon J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guatin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragae V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051775; AAK93199.1; -; mRNA.
DR HSSP; O67082; 1JF2.
DR Ensembl; CG8730; Drosophila melanogaster.
DR Flybase; FBgn0026722; CG8730.
DR Flybase; FBgn0026722; drosha.
DR GO; GO:0005634; C.nucleus; IEA.
DR GO; GO:0003725; F.double-stranded RNA binding; IEA.
DR GO; GO:0004519; F.endonuclease activity; IEA.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR GO; GO:0004525; F.ribonuclease III activity; IEA.
DR GO; GO:0006396; P.RNA processing; IEA.
DR InterPro; IPR001159; D5_RBD.
DR InterPro; IPR000999; RNasec_III.
DR Pfam; PF00035; darm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS00137; D5_RBD; 1.
DR PROSITE; PS00517; RNasec_3; 1.
DR PROSITE; PS00142; RNasec_3; 2.
DR PROSITE; PS00142; RNasec_3; 2.
SQ SEQUENCE 1071 AA; 123158 MW; 25E6C2AADA5078B CRC64;

Query Match 37.5%; Score 2811.5; DB 2; Length 1071;
Best Local Similarity 54.6%; Pred. No. 5.3e-129;
Matches 545; Conservative 169; Mismatches 220; Indels 65; Gaps 15;

QY 412 WIRCTHSENYSSDPMDQVGDSTVGTSTRLRLDYKFEELSGROKAKAARPEWEPKPT 471
DB 14 WVRSSPALEYERTK---SENEVRGRARLQKLTLPDEBLQRAKRVREKLPVYVPPR 69
QY 472 KL-----DEDESSSESCEDESDSTGSSSSDSSEFV-----IAETKKAAND 516
DB 70 KARRRVCKHKHSEACSSSSSSDDDEDA-----FKLEQDCMEELSRKVQHQ 119
QY 517 RLHDELWYNDPGQMGDPCKCKAKARTGIRHSIYGEERAPRCPTMNAARLPHYRI 576
DB 120 RYHADLMHNDAGMGNDGPLCRCKAKSRIRGIRGITPGEIYGLCDPNSNNAKGLHYRI 179
QY 577 TVSPPTNPLDRPTVLEYDDHEVTFEGFSMFAPLPTNPLCKVIRFNIDYTIHFIEMK 636
DB 180 SISPPNPLTKPTIILKHDEHFLFSGFSLSHVRLSDLPVCVIFNIEYTIYEDEKX 239
QY 637 PENFCVKGLELSEFLFRDILLEYDMNLKGPL-----FEDSPCCPFPHMPRVRLPDS 692
DB 240 PENFTIHELDIFPKYLFHELELVDFNLMPNLPSGNVEES---CPAFHFEPFRVRLPDR 296
QY 693 GKEVLSMHOILYLLACSKALVPEEBRLANLQWELMOKVYAECKGMTYTPGTPSSV 752
DB 297 GKEVLAMVEELRLKLNLSAQLVERQQLHLNQISQSEWQYVDFPIKMLTKRKGYPCSI 356
QY 753 RIDQDREOFNPDV-----ITPFIIVHGIRPAOLSYAGDPQYOKLMKSYVXLRLH 803
```

```
DB 357 RVDQDRLN--NSDLBECVDRETGISHPAIVHGIQHPQSLSTAGNBYQAMERYKYRHL 414
QY 804 IANSKRVQOTDKOKLAOREEALOKIROKNTMREYTVELSSQGFMTGIRSDVCQHAMML 863
DB 415 MANNKSPFKDKRKLKEKQRLQEMRTQGRMBRNITVAISSGFRITGIMCVQJAMLI 474
QY 864 PVLTHIRYHOCMLMDKILIGTFODRCLQLAMTHPSHHNLFGANPDHARSLSNGIR 923
DB 475 PVLTHIRYHOCMLMDKILIGTFODRCLQLAMTHPSHHNLFGANPDHARSLSNGIR 923
QY 924 QPKYGDRLVHMHMKKGINTLINIMSRGODDPTPSRINHERLEFLDAAVEFLTSYH 983
DB 535 QPEYGDRLKIHNTKRGINTLVSTMSRFGKHEVTSNITHERLEFLDAAVEFLTSYH 594
QY 984 LYYLPPSLEEGGLATYRTAIVONQHILAMAKKLEDPFLYAHGPDLCRESDLRHAMANC 1043
DB 595 LFFMPELEEGGLATYRTAIVONQHILAMAKKLEDPFLYAHGPDLCRESDLRHAMANC 654
QY 1044 FEALIGAVYLEGSLBEAKQFGRLLP--NDPDLREAVWLYPLHPLQLEPNTROLITSP 1102
DB 655 FEALIGAVYLEGSLBEAKQFGRLLP--NDPDLREAVWLYPLHPLQLEPNTROLITSP 1102
QY 1103 VLQKITEPBEALGVFTYVRLRLARAFTLRTVGFNHLTGHNORMEFLGDSIMQVATEYL 1162
DB 715 VLKELTKREDSIGIFKHLRLARAFTDRSIGFTHLTLGSSNORLEFLGDTVLQICSBYL 774
QY 1163 FIFHPDHHEGHTLLRSSLVNNRTOAKVAEELGMOBYAI--TYDKTKRPVGLTKTLADLL 1221
DB 775 YHFFPEHHEGHTLLRSSLVNNRTOAKVAEELGMOBYAI--TYDKTKRPVGLTKTLADLL 1221
QY 1222 ESFIALYTDKLEVYHTFMVYCFPRLKEFLINQDMDNPKSLOQCCITLRT--EGKEPD 1280
DB 831 EAFGLALYVDKGLTYCEGCHVCLFPRQLQFLINQDMDNPKSKLOQCCITLRTMDGEBD 890
QY 1281 IFLYKTLQVPSHARTTVAYYFGERIGCGCKGSIQQAEGAMDALEKXN--PPOVA 1338
DB 891 IFYKTVASGPTNRYKVAAYFRSKRLATSSGSSSIQQAENNAAKALENSRDLPPQD 950
QY 1339 HOKRFIGRKYROEL---KEMRWEREHOEREPDETDIKK 1374
DB 951 HOKRVIATKIKQGTGNELDNDSRQHOE-----EKIKR 983

RESULT 14
Q7PNE9 ANOGA
ID Q7PNE9 ANOGA PRELIMINARY; PRT; 836 AA.
AC Q7PNE9;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000011622 (Fragment).
GN ORFNames=ENSANGG0000009133;
OS Anopheles gambiae str. PE8T.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_Taxid=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PE8T;
RA The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PE8T;
RA The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008964; BAA12298.3; -; Genomic_DNA.
```

DR GO:0005622; C:intracellular; IEA.
DR GO:0003725; F:double-stranded RNA binding; IEA.
DR GO:0004519; F:endonuclease activity; IEA.
DR GO:0004525; F:ribonuclease III activity; IEA.
DR GO:0006396; F:RNA processing; IEA.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dnm; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS0137; DS_RBD; 1.
DR PROSITE; PS00517; RNase_3_1; 1.
DR PROSITE; PS0142; RNase_3_2; 2.
FT NON TER 1 1
FT NON TER 836 836
SQ SEQUENCE 836 AA; 96229 MW; 8D3E0BD04953877D CRC64;

Query Match 35.8%; Score 2683.5; DB 2; Length 836;
Best Local Similarity 59.2%; Pred. No. 76-123;
Matches 499; Conservative 143; Mismatches 182; Indels 19; Gaps 8;

QY 514 HPRLHDELMTYNDPGQNDGRLCKSAKARTGIRHSITYPEGALIRCPMTNAGRLTH 573
DB 1 HPRLHVDLWYNDPGEMNDGRLCRCSARSRTGIRHGKYPGEBGPKCIPNSNNAKLTH 60
QY 574 YRITVSPPTNPLDRPTVIEYDDEHYIPGFSMFAPHAPLTNPLCKVIRPNIDYTHFTF 633
DB 61 YRITVSPPTNPLKTPTIIGHDHEPLRBFSLAHEPLIGELTCKVIRNIEYTHLYH 120
QY 634 EMMPEFCVKGLELPSLFRDILIELYDMLKGPLFEDSPCCRFHMFPRFVRLPDG 693
DB 121 EQMPEFTIRBELNLFRIYFRELELDVDTVQ-PSGSGESSGCPVHFLPRFRDLPDG 179
QY 694 KEVLSMHQILLYLRCSKALVPEEELANMLQWEELEWOKYAECKMITYNPTKSSVR 753
DB 180 KEVLSMHQILLYLRCSKALVPEEELANMLQWEELEWOKYAECKMITYNPTKSSVR 239
QY 754 IGDLEGGQNDPTPTIIVHGIIPAQLSYAGDPQYKMSVYVLRHLLANSPLVKQT 813
DB 240 VQGLDNNV--GDVPEANALIDENG-----YAGNPEYQAKREYIIFRHLIAMSXPSE 290
QY 814 DRQKLAQREBALOKIRQKNTMRREVVELSSQGFMTGIRSDVCOHAMLPLVTHHRYH 873
DB 291 DRKLEAKENRLLLEMMQGMKNITITAVSAKAFHGTGIMCDVQDAMLPLVTHHRYH 350
QY 874 QCLMHLDKLIGYTFQDRCLQLAMTHPSHHNLFQGNPDHARNSLSCGIRQPKYGRKYH 933
DB 351 RSLNVLERYIGYFTNRVTLQLALTHPSYKENGFTNPDHARNSLSCGIRQPKYGRKI 410
QY 934 HHMRKKGINTLINISRLGODDPTPSRIINHERLEFLGDVAVVEFLTSVHLVYLPFSL 993
DB 411 YNMTRRGINTLISINSRFGKEHETSNTTHNERLEFLGDVAVVEFLTSVHLVYLPFSL 470
QY 994 GGLATYRTAIVONQHAMLAKKLELDPFMLYAGPLCESDRLHMANCFEALIGAVY 1053
DB 471 GGLATYRTAIVONQHAMLAKKLELDPFMLYAGPLCESDRLHMANCFEALIGAVY 530
QY 1054 ESSLERAKQLFGRLLFNDDP-LREVMNLNPLRLQLQEPNTDRLISTSVLOKTEPFE 1112
DB 531 DGGIEVADRVFAVALFOEDDTLRGIWVNPSPHLQOEPLGDRHNLDSFEMLKTLPFED 590
QY 1113 AIGVITTHRLARATLRTVGNHLLTGNNQMEFLGDSIMQVATVYLFTHPRHHG 1172
DB 591 SIVGVNHNRLRLARATLRTVGNHLLTGNNQMEFLGDSIMQVATVYLFTHPRHHG 650
QY 1173 HLLTSSSLVNNRQAQVAELGMOEYAI-TNDKTRPVGLRTKTLADLLESFIALYTD 1231
DB 651 HLLTSSSLVNNRQAQVAELGMOEYAI-TNDKTRPVGLRTKTLADLLESFIALYTD 706
QY 1232 KDLVYVTHMNVCFPRLEKEFLINQDNDPKSQLQCCCLTART-ESEKPDIPLYKTLQTV 1290
DB 707 KGLVYCEMCHVCLFPRLODFINQDNDPKSKLQCCCLTARTMDGEPDIPYKYIECT 766

QY 1291 GSHARTYVAVYFKERIGCGKPSIQQAENGAMDALE--KYNPPQAHQKRTIGRX 1348
DB 767 GPTNTRVAVAVYFRKRLACADGHSIQQAENMAAQAQLENSDLPQLDHQGRVIAQSL 826
QY 1349 RQE 1351
DB 827 KRQ 829

RESULT 15

QY 059FP9 HUMAN PRT; 432 AA.
ID 059FP9_HUMAN PRELIMINARY;
AC 059FP9;
DT 10-MAY-2005 (TEMBLrel. 30, Created)
DT 10-MAY-2005 (TEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TEMBLrel. 30, Last annotation update)
DE Nuclear RNase III Drosha variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209501; BAD92738.1; -; mRNA.
FT NON TER 1 1
SQ SEQUENCE 432 AA; 50659 MW; E5D81B6DD61251A CRC64;

Query Match 30.6%; Score 2298; DB 2; Length 432;
Best Local Similarity 99.8%; Pred. No. 2,1e-104;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 556 BAIKPCRPMTNNAAGRLFHYRITVSPPTNPLDRPTVIEYDDEHYIPGFSMFAPHAPLTN 615
DB 7 QAIKPCRPMTNNAAGRLFHYRITVSPPTNPLDRPTVIEYDDEHYIPGFSMFAPHAPLTN 66
QY 616 PLCKVIRFNIDYTHIRIEEMMPENFCVKGLELPSLFRDILIELYDMLKGPLFEDSPC 675
DB 67 PLCKVIRFNIDYTHIRIEEMMPENFCVKGLELPSLFRDILIELYDMLKGPLFEDSPC 126
QY 676 CPEFHMPFRVRLPDGKREVSMDQILLYLRCSKALVPEEELANMLQWEELEWOKYAE 735
DB 127 CPEFHMPFRVRLPDGKREVSMDQILLYLRCSKALVPEEELANMLQWEELEWOKYAE 186
QY 736 ECKGMIVTNPGRKPSVRIDOLDRQFNPDVTFPIIVHGIIPAQLSYAGDPQYKLMK 795
DB 187 ECKGMIVTNPGRKPSVRIDOLDRQFNPDVTFPIIVHGIIPAQLSYAGDPQYKLMK 246
QY 796 SYVKLHLLANSPLKVQTDKQKLAQREBALOKIRQKNTMRREVVELSSQGFMTGIRSD 855
DB 247 SYVKLHLLANSPLKVQTDKQKLAQREBALOKIRQKNTMRREVVELSSQGFMTGIRSD 306
QY 856 VCOHAMLPLVTHHRIYHOCMLHDLKLGITFQDRCLQLAMTHPSHHNLFQGNPDHAR 915
DB 307 VCOHAMLPLVTHHRIYHOCMLHDLKLGITFQDRCLQLAMTHPSHHNLFQGNPDHAR 366
QY 916 SLNSCGIRQPKYGRKRVHMHMRKKGINTLINMSLGGDDPTPSIINHERLEFLGDV 975
DB 367 SLNSCGIRQPKYGRKRVHMHMRKKGINTLINMSLGGDDPTPSIINHERLEFLGDV 426
QY 976 VERLT 980
DB 427 VERLT 431

Search completed: February 3, 2006, 19:41:12
Job time : 297 secs

```

QY 421 YSSDPMDOVGDSTVVGTSRLRLDYDKFEEELSGROEKAARPMWPKTKLDEBLESS 480
|
|
|
Db 421 YSSDPMDOVGDSTVVGTSRLRLDYDKFEEELSGROEKAARPMWPKTKLDEBLESS 480
QY 481 SESECESDSDSTCSSSDSEVPDYIAEIRKKAHPRLHDELMYNDPGQNDGPLCKCSA 540
|
|
|
Db 481 SESECESDSDSTCSSSDSEVPDYIAEIRKKAHPRLHDELMYNDPGQNDGPLCKCSA 540
QY 541 KARRGIRRSIYRGEBAIKPCRPMTNNAARLPHYRTTSPNMFILDRPTVIEYDDEHY 600
|
|
|
Db 541 KARRGIRRSIYRGEBAIKPCRPMTNNAARLPHYRTTSPNMFILDRPTVIEYDDEHY 600
QY 601 FEGFSMFAHPLTNILPLCKVIRFNIDYTHFIEEMAPENFCVGLFSLFLFRDILEY 660
|
|
|
Db 601 FEGFSMFAHPLTNILPLCKVIRFNIDYTHFIEEMAPENFCVGLFSLFLFRDILEY 660
QY 661 DNMLKGPLFEDSPCCPRFHFMRFRVRLPDGKEVLSMHQILLYLRCSKALVPEEBA 720
|
|
|
Db 661 DNMLKGPLFEDSPCCPRFHFMRFRVRLPDGKEVLSMHQILLYLRCSKALVPEEBA 720
QY 721 NMLQWEELEWQYAECECKMIVTNPCTKPSVRIIDQDRBQFNPDIYTPILVHGIKRA 780
|
|
|
Db 721 NMLQWEELEWQYAECECKMIVTNPCTKPSVRIIDQDRBQFNPDIYTPILVHGIKRA 780
QY 781 QLSYAGDPQYQKLWKSUYVLRHLANSPKVKQTKQKLAQREBAOKIRQKMTMREEVY 840
|
|
|
Db 781 QLSYAGDPQYQKLWKSUYVLRHLANSPKVKQTKQKLAQREBAOKIRQKMTMREEVY 840
QY 841 ELSSQGFMTKGTIRSDVCQHAMMLPVLTNHIRYHQCIMHLDKLGTFORCLQLAMTBP 900
|
|
|
Db 841 ELSSQGFMTKGTIRSDVCQHAMMLPVLTNHIRYHQCIMHLDKLGTFORCLQLAMTBP 900
QY 901 SHHLNGMNPDBHARNSLSCGTRQPKYGRKVMHMRKKGINTLINMSRLGODDPTSS 960
|
|
|
Db 901 SHHLNGMNPDBHARNSLSCGTRQPKYGRKVMHMRKKGINTLINMSRLGODDPTSS 960
QY 961 RINNNRLEFLGDVAVEFLTSVHLVYLPSSLBEGGLATYRTAIVQNHMLAKKLEDP 1020
|
|
|
Db 961 RINNNRLEFLGDVAVEFLTSVHLVYLPSSLBEGGLATYRTAIVQNHMLAKKLEDP 1020
QY 1021 FMLYAGPRLCRESDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMIN 1080
|
|
|
Db 1021 FMLYAGPRLCRESDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMIN 1080
QY 1081 YPLHNPQLQDEPNDRQLIENSPVLOKLTBEFEBAIGVIFTHVRLARAFTLRGVFNHLLT 1140
|
|
|
Db 1081 YPLHNPQLQDEPNDRQLIENSPVLOKLTBEFEBAIGVIFTHVRLARAFTLRGVFNHLLT 1140
QY 1141 GHNQRMFLGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYA 1200
|
|
|
Db 1141 GHNQRMFLGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYA 1200
QY 1201 ITNDKTRPVGLRTKTLADLLESFIALYTDKLEVATFMNVCFFPRLKEFLINDQMD 1260
|
|
|
Db 1201 ITNDKTRPVGLRTKTLADLLESFIALYTDKLEVATFMNVCFFPRLKEFLINDQMD 1260
QY 1261 PKSQLOOCCCLTATEKEPDIPLKYTLQTVGSHARTYTAAYVFKSGRIGCGSPSIQQA 1320
|
|
|
Db 1261 PKSQLOOCCCLTATEKEPDIPLKYTLQTVGSHARTYTAAYVFKSGRIGCGSPSIQQA 1320
QY 1321 EMGAAMDALEKYNPQMAHQKRFIGRKYRQELKEMWEREHOBERPDETDIKK 1374
|
|
|
Db 1321 EMGAAMDALEKYNPQMAHQKRFIGRKYRQELKEMWEREHOBERPDETDIKK 1374

```

```

RESULT 2
US-09-900-425B-37
; Sequence 37, Application US/09900425B
; Patent No. 6737512
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.

```

```

; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425B
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 466
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-900-425B-37

```

```

Query Match 32.8%; Score 2463; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 2,4e-185; Indels 0; Gaps 0;
Matches 466; Conservative 0; Mismatches 0;

```

```

QY 909 NPDHARNSLSCGTRQPKYGRKVMHMRKKGINTLINMSRLGODDPTSPRINNEBL 968
|
|
|
Db 1 NPDHARNSLSCGTRQPKYGRKVMHMRKKGINTLINMSRLGODDPTSPRINNEBL 60
QY 969 EFLGDVAVEFLTSVHLVYLPSSLBEGGLATYRTAIVQNHMLAKKLEDPMLYAGP 1028
|
|
|
Db 61 EFLGDVAVEFLTSVHLVYLPSSLBEGGLATYRTAIVQNHMLAKKLEDPMLYAGP 120
QY 1029 DLCRESDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMINPLHPLQ 1088
|
|
|
Db 121 DLCRESDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMINPLHPLQ 180
QY 1089 QEPNDRQLIENSPVLOKLTBEFEBAIGVIFTHVRLARAFTLRGVFNHLLTGHNORMEF 1148
|
|
|
Db 181 QEPNDRQLIENSPVLOKLTBEFEBAIGVIFTHVRLARAFTLRGVFNHLLTGHNORMEF 240
QY 1149 LGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTKR 1208
|
|
|
Db 241 LGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTKR 300
QY 1209 PVGLRTKTLADLLESFIALYTDKLEVATFMNVCFFPRLKEFLINDQMDPKSQLOOC 1268
|
|
|
Db 301 PVGLRTKTLADLLESFIALYTDKLEVATFMNVCFFPRLKEFLINDQMDPKSQLOOC 360
QY 1269 CCTLRTGKEPDIPLKYTLQTVGSHARTYTAAYVFKSGRIGCGSPSIQQAEMGAAMDA 1328
|
|
|
Db 361 CCTLRTGKEPDIPLKYTLQTVGSHARTYTAAYVFKSGRIGCGSPSIQQAEMGAAMDA 420
QY 1329 LEKYNPQMAHQKRFIGRKYRQELKEMWEREHOBERPDETDIKK 1374
|
|
|
Db 421 LEKYNPQMAHQKRFIGRKYRQELKEMWEREHOBERPDETDIKK 466

```

```

RESULT 3
US-09-900-425B-3
; Sequence 3, Application US/09900425B
; Patent No. 6737512
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425B
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 412
; TYPE: PRF
; ORGANISM: Caenorhabditis elegans
US-09-900-425B-3

```

```

Query Match 11.9%; Score 890; DB 2; Length 412;
Best Local Similarity 43.1%; Pred. No. 1,4e-61;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;
QY 944 TLINMSRLGODDPTSPRINNNRLEFLGDVAVEFLTSVHLVYLPSSLBEGGLATYRTAI 1003

```

```
Db      2 SLFNKGTSGGSP-----ILHNERLEYGDVAVVELIVSHLTYMLTHHFEFGIATYRTAL 57
Qy      1004 VONOHAMLAIAKKLELDPFMLYAHGPDLCRESDLRHMANCFBALIGAVYLEGSLBEAKOL 1063
Db      58 VONRNATLAKNCRIEMLOYSHGADILINAEFKHALANAEVMAIYIDGGLACDVI 117
Qy      1064 FGLFLF-NDPDLREVLNYPPLHPLLOEPNTDROLIETSPVLOKTEFEBAIVIFTHVR 1122
Db      118 FSAWYGHQVLEKKMDHINEHELKREDPGDRDLSEITPTLTSPHAEERLDIGFNINR 177
Qy      1123 LARATLTGNNHLLTGNOMERPLGDSIMOLVATEYVPIHPDHHSGLTLLSSLY 1182
Db      178 LAKAFTERRRIPNNDLTGHNQRLWLGSLVLOLIVSDFLYRRFPYHNEGHMSLRTSLV 237
Qy      1183 NNRTOKVAEELGMOEYATINDTKRPV---GLRTLDLLESFIALYTDLDLEVHT 1239
Db      238 SNOTQAVCCDGLFTEFVI-----KAPYKTPBLKLDKADLVFAFGALYVDRGIEHCRA 292
Qy      1240 FMNVCFEPPLKEFILNQDNDPKSOLOCCCLTLR-TEGKEPDIPLYKTLQTVGSPHARTY 1298
Db      293 FIRIVCFPLKHIIESEKNDAKSHLOQKCLAMRDSSSPDMPEYRVLGIEGPTNNRIF 352
Qy      1299 TAAVYFKERIGCGKPSIQOAEKGA---AMDLEKYNPOM-AHOKRPIGRYQELKE 1354
Db      353 KIAVYKGRKLASAAESNVKAEKLVAEELALANLLESMFSKAKANNMNRRLLEDQDTS 412
```

RESULT 4

```
US-09-513-999C-4635
; Sequence 4635, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclercq, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4635
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23..-1
; OTHER INFORMATION: score 5.9
; OTHER INFORMATION: seq TLKFLTLLOKSNM/KR
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 3
; OTHER INFORMATION: Xaa=Ala or Pro
US-09-513-999C-4635
```

```
Query Match      6.6%; Score 496; DB 2; Length 115;
Best Local Similarity 95.7%; Pred. No. 2, 2e-31;
Matches 88; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      511 KKAHPRLDHDELYNDPGQNDPGLCKCSAKARTGIRHSIYGEBAIKRCRPMYNAGR 570
Db      24 KRANLRLHDELYNDPGQNDPGLCKCSAKARTGIRHSIYGEBAIKRCRPMYNAGR 83
Qy      571 LPHYRTTSPPTNFLTDRPTVIEYDDHEYLFE 602
Db      84 LPHYRTTSPPTNFLTDRPTVIEYDDHEYLFE 115
```

```
RESULT 5
US-09-949-016-7961
; Sequence 7961, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7961
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7961
```

```
Query Match      4.2%; Score 316; DB 2; Length 553;
Best Local Similarity 28.0%; Pred. No. 3, 2e-16;
Matches 111; Conservative 32; Mismatches 106; Indels 148; Gaps 21;
```

```
Qy      19 PRGRGHGAPSPASFRPONLRLHPOQPPVOYQYPPSPASTTFSNSPAPNF---LPP 74
Db      200 PQGGRGRGPRPGA-----VFGGDRFPGPAGP---CGPPPPPPAGQTP 239
Qy      75 RPDVFPFPPMPSPAGPLPBCIRPP-----FBNHGMHHPFY-----PPC---PPP 119
Db      240 RP---PLGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 236
Qy      120 MPPMPGCPNNPYPVGAAPGQ-----GTFP-----FMMPSPMHPPP---PPV 159
Db      297 GPPGPP-FVPGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 355
Qy      160 MPOQVNYQYPPGYSHNPPPPSPFNQNNPSPFLPANNSSSP---HFRHLPY----- 210
Db      356 AP-HVN---PAF---PPPT-----NSGMPYSDRGPPTDPYGRPPYDRGDYG 398
Qy      211 PLPKASERSRSLKHYDDHRHDS----- 237
Db      399 PPGREMDTARTPLSEAFEFIMNRRAISSAISRAVSDASAGDYSAIETLVTAISLIK 458
Qy      238 -----HG-RGERHRLDRRGRSPDRRQDSRYSDYDRGRT 275
Db      459 QSRVSAADRCVKYLISLQDCLHGISKSYGSGRRSRERDRHSR---SRKSKRRHKSRSR 516
Qy      276 SRHSYRSRERERERHRHNRSPSLRSYKEXK 312
Db      517 DRHDYRERSRERERHRDRD--RDRDRERDRERER 551
```

RESULT 6

```
US-09-107-532A-5434
; Sequence 5434, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
```

```
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
```

```

1      ZIP: 02354
2      COMPUTER READABLE FORM:
3      MEDIUM TYPE: CD-ROM ISO9660
4      COMPUTER: PC
5      OPERATING SYSTEM: <Unknown>
6      SOFTWARE: ASCII
7      CURRENT APPLICATION DATA:
8      APPLICATION NUMBER: US/09/107,532A
9      FILING DATE: 30-Jun-1998
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER: 60/085,598
12     FILING DATE: 14 May 1998
13     APPLICATION NUMBER: 60/051571
14     FILING DATE: July 2, 1997
15     ATTORNEY/AGENT INFORMATION:
16     NAME: Arinello, Pamela Deneke
17     REGISTRATION NUMBER: 40,489
18     REFERENCE/DOCKET NUMBER: GTC-012
19     TELECOMMUNICATION INFORMATION:
20     TELEPHONE: (781)893-5007
21     TELEFAX: (781)893-8277
22     INFORMATION FOR SEQ ID NO: 5434:
23     SEQUENCE CHARACTERISTICS:
24     LENGTH: 241 amino acids
25     TYPE: amino acid
26     TOPOLOGY: linear
27     MOLECULE TYPE: protein
28     HYPOTHEetical: YES
29     ORIGINAL SOURCE:
30     ORGANISM: Enterococcus faecium
31     FEATURE:
32     NAME/KEY: misc feature
33     LOCATION: (B) LOCATION 1...241
34     SEQUENCE DESCRIPTION: SEQ ID NO: 5434:
35     US-09-107-532A-5434

```

```

Query Match 4.2%; Score 312.5; DB 2; Length 241;
Beet Local Similarity 33.5%; Pred. No. 1,8e-16;
Matches 77; Conservative 50; Mismatches 90; Indels 13; Gaps 7;

QY 1105 OKLTFFEEAIGVIFTHVRLAARFLLRTVGFNH--LTIGHNORMEFLGDSIMQVATEYL 1162
DB 17 QLTREKKNENYIVFDFDNLBLEQAFTHSSVYVNEHRNQLSDNRELFELGDAVLELWVSQYL 76
QY 1163 FIFHPDHHEGHTTLRRSLVANNRTOAKVAEEIGMOEYALTNDKTRPVGLRTK--TLADL 1220
DB 77 FRLVPHLPEGGKLTIKTKRAAIVREDSLSKFAKXCHFPQYVLLG--KGEENSGGRTRPALLCPL 135
QY 1221 LESFPAALYTPOKDLLEVHTFMNVCFEPPRLKEFTLINDDNMDPKSLOQCCCLTRTEGKEPD 1280
DB 136 FEAFAGALYLDQGFYTTAFLEKVIFFPKVKAGAFSHEM-DHKTKLOE---VLQXSG--D 188
QY 1281 IPL-YKTLQTVGSPSHARTYTVAVVYFKGERIGCGKSPISQOAEEMGAMDAL 1329
DB 189 VSIEFRLINEGPAHERVFWIEVYDDQLIGTGCKSKKLAAQAAENNL 238

RESULT 7
US-09-583-110-4554
Sequence 4554, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
FILE REFERENCE: PATHO0-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553

```

```

; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4554
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-563-110-4554

Query Match      3.8%; Score 284.5; DB 2; Length 232;
Best Local Similarity 32.9%; Pred. No. 2,86-14;
Matches 76; Conservative 48; Mismatches 90; Indels 17; Gaps 8;

QY   1108 TEFEBAIGVIFTHVRLLLARAFTLTGVGNH--LTLGHNQRMEFGDSIMOLVATEYLFIH 1165
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB    6  TLANHFPAIEPTDKCLLETAFTHTSYANHERHLLKISHNERLEFPGDAVLQLLSIERYLYKK 65

QY   1166 FPDHHGEHLTLRSLSVNNRTQAQVAEELGMOEAIVINDKTGRVGLRTK--TLADLLES 1223
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| :
DB    66 YPKREPGDLSTLRAMIVREESLAGFARDCOFPDF-IKLKGKEEKSGGNNRDTIIIGDAFEA 124

QY   1224 FIALLTYDIXOLEVYHTFMNVCFPRRK--FFILNQDNNDPSSQLQQOCCVLTLTREGKEPDI 1281
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB    125 FLGLALLDKDQAKAKEFIYQMIPKVAEGEFEMITDY---KTHIQE--LIQVNG---DV 175

QY   1282 PL-KETLQTQVPSHARTYTVAVVYFKGRIQGGKGPSICQAEMLGAMDALXK 1331
      ||| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB    176 AIRQVTSISETPADIKVFDEVEVLVEGSIIGGGORSRSKKLAEQAANAVERK 226

```

```

1  RESULT 8
2  US-09-107-433-4394
3  : Sequence 4394, Application US/09107433
4  : Patent No. 6800744
5  : GENERAL INFORMATION:
6  : APPLICANT: Lynn A Doucette-Stamm and David Bush
7  : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
8  : SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS
9  : THERAPEUTICS
10 :
11 : NUMBER OF SEQUENCES: 5206
12 : CORRESPONDENCE ADDRESS:
13 : ADDRESSEE: GENOME THERAPEUTICS CORPORATION
14 : STREET: 100 Beaver Street
15 : CITY: Waltham
16 : STATE: Massachusetts
17 : COUNTRY: USA
18 : ZIP: 02354
19 :
20 : COMPUTER READABLE FORM:
21 : MEDIUM TYPE: CD-ROM ISO9660
22 : COMPUTER: <Unknown>
23 : OPERATING SYSTEM: <Unknown>
24 : SOFTWARE: <Unknown>
25 :
26 : CURRENT APPLICATION DATA:
27 : APPLICATION NUMBER: US/09/107,433
28 : FILING DATE: 30-Jun-1998
29 :
30 : PRIOR APPLICATION DATA:
31 : APPLICATION NUMBER: 60/085131
32 : FILING DATE: May 12, 1998
33 : APPLICATION NUMBER: 60/051553
34 : FILING DATE: July 2, 1997
35 :
36 : ATTORNEY/AGENT INFORMATION:
37 : NAME: Arinello, Pamela Deneke
38 : REGISTRATION NUMBER: 40,489
39 : REFERENCE/DOCKET NUMBER: GTC-011
40 : TELECOMMUNICATION INFORMATION:
41 : TELEPHONE: (781)893-5007
42 : TELEFAX: (781)893-8277
43 :
44 : INFORMATION FOR SEQ ID NO: 4394:
45 : SEQUENCE CHARACTERISTICS:
46 : LENGTH: 236 amino acids
47 : TYPE: amino acid
48 : TOPOLOGY: linear
49 : MOLECULE TYPE: protein
50 : HYPOTHEetical: YES
51 :

```


ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: m1ec feature
LOCATION: (B) LOCATION 1...236
SEQUENCE DESCRIPTION: SEQ ID NO: 4394;
US-09-107-433-4394

Query Match 3.8%; Score 284.5; DB 2; Length 236;
Best Local Similarity 32.7%; Pred. No. 2.8e-14;
Matches 76; Conservative 48; Mismatches 90; Indels 17; Gaps 8;

QY 1108 TERBEAIGVIFTHVRLAARFLRTVGPNH--LTGHNGMEPLGDSIMQVATEYLFIH 1165
DB 10 TVLKNHFALEFDTKLETAFTHTSYANERHLKISHNERLEFLGAVQLLSELYKK 69
QY 1166 PPDHGHGLTLRSLVNNRTOAKVABELOEYAITNDKTRPVGLRTK--TLADLLES 1223
DB 70 YPKKPEGDLISKRAMIVRESLAGFARDQCPDQF--IKLGKBEKSGGRNDTLIGDAFEA 128
QY 1224 FIAALYTDLDLVYHFMVAVCFPPRLK--EFLINDMDPKSQLOCCULTLRTGKEPDI 1281
DB 129 FIALALLDDVAKVKEFTYQVMIPKVEAGEFEMITDY--KTHLQE--LLQVNG--DV 179
QY 1282 PL-YKTLQTVGSPHARTYVAVYFKGERIGCGKPSIOQAEMGAMDALEK 1331
DB 180 AIRYQVISSTGPAHDKVFDVEVLVEKSGIGGGRKLAEGEAKNAVEX 230

RESULT 9
US-09-949-016-7757
Sequence 7757, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7757
LENGTH: 998
TYPE: PRT
ORGANISM: Human
US-09-949-016-7757

Query Match 3.8%; Score 284.5; DB 2; Length 998;
Best Local Similarity 22.7%; Pred. No. 2.3e-13;
Matches 151; Conservative 63; Mismatches 206; Indels 245; Gaps 30;

QY 27 ARPSA--PSFRPQNLRLHQPQPPVQYQYPPSA-----PSFTFSNPAFLPR 75
DB 350 ARGGGRPGWEHNS-KLGYLVSPQQRGRGSCYSINNGRHSRSQRTQPS--LPAT 407
QY 76 PPDVPPRPMPRAQGRPLRPPCRPRPPNQHMRPPVPCFPMPMPMCPNNPVPGA 135
DB 408 PVFVVPVPP-----PLVPPPP--HTLPLP--GCVPPQPSQFP--PQO 445
QY 136 PPGQTFPPMPPSPMHPRPVPMQOVN-----YQYPP-----170
DB 446 PPRAG--YVPPRPGP--PAPANLSTPWSGVTAAHSNTIPTTQAPRLSREFFYEQRR 501
QY 171 -----GYSHNFPSPSFPNS 184
DB 502 LKEEKKKSLDEFTNDAKELMEYKKIQKERRRSFSRSKSPYSGSSYSTYSSKRS 561

QY 185 PQNNPSPFLPSANSSSPHFRHLPPYPLPKAPSERSPERLKYDDHRHDSHGGERH 244
DB 562 GSTRSRSYRSRSPSRSHRSRYSNRPYP-----RRGRGSRNY--RGRSRHG--YH 608
QY 245 RSLDRRERKGRSPDRRQDSRYRSDYD--RGRTPSRHSYERSRERER-----289
DB 609 RS-----RSRSPYRRYHSRSRSPQAFRCQSPPKRVAVPQGETREYFNRYREVPYDMK 663
QY 290 -----ERRHRDNRSPLERSYKKEVRSRGSVLSVVPAGCTPELPGE 336
DB 664 AYYGRSVDPDPPEKERYE-----WEKRYREWEKYKGYAAGQAPRSA-----709
QY 337 IINKTDSMAP-----PLEIVNHRSPSRREKTRARWESEKDWSDQSSG-----KD 381
DB 710 ---NRNFSRPERLPANTIN--SPFRGRREDYVGQSHRSRNIGSNYPEKLSABDGHQ 764
QY 382 KNYTSIKEKPEFTMD-----KNEBEBELLKPWIRCTH-----SENY 421
DB 765 KONTKSKEKSENAVAPDQGNKHKHKKRRKRGSESEGFNPBELLTSRKSREPTVEENK 824
QY 422 YSS-----DPMQVGDSTVGTSLRLDLYDKPEEL-----GSRQ 456
DB 825 TDSLFLPBRDDATPYRDEPMD--AESITPKSVSEKQKREDPKAKGDKTKRKNDGSAY 882
QY 457 EKAKARPWEPPTKTLDEDLSESSSESEC-----ESDEDSTCSSSDSEVPDVIARIK 509
DB 883 SKKENVYKAKGQEKVDGERERSPPRPKAKKETKTDTNTKSSSSSQDKEKITGTP 942

QY 510 RKXAH 514
DB 943 R-KAH 946

RESULT 10
US-08-869-674-2
Sequence 2, Application US/08869674
Patent No. 5866365

GENERAL INFORMATION:
APPLICANT: Lonetto, Michael A.
APPLICANT: Rosendberg, Martin
TITLE OF INVENTION: NOVEL COMPOUNDS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/869,674
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: GM10013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,308
FILING DATE: 18-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 372,643P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-308-1

Query Match 3.5%; Score 264.5; DB 2; Length 1404;
Best Local Similarity 19.9%; Pred. No. 1,4e-11;
Matches 214; Conservative 125; Mismatches 327; Indels 411; Gaps 50;

QY 25 HGARPSA--PSRPQNLRLHPQOPVQYQYEPSPA-----BSTFSNSPAPNPLP 73
DB 314 NARPCGGRPGWHSN-KLGYLVSPQQLRGRSCYRSINRGHHSEKQTSQPS-LP 371
QY 74 PRPDVFPFPPPPPSAQGLPFCPIRPPPNHQMHPFVPPPCFPMPMPMCPNNPPVP 133
DB 372 ATPCFVPVPP-----PPLYPFPHTLPLP-----P 396
QY 134 GAPPGGT--FPPMPSPMPHPPPVMPQVNYQ--YPPGY--SHNFPSPS----- 181
DB 397 GVPPOFSQFPSSQPTAGYSVPPGFPAPANISTACSPGVPTAHSTMTPTQAPLL 456
QY 182 ----FNSFQNP-----SSFPSANNS-----SPHFRH---LPPY 210
DB 457 SREEFREGNDKGRSKFPYSSSYSSSYTDSQGLAQHIALTSPSAHTLDLHDH 516
QY 211 PLPKASERSRRLGHYDHRHDHSHGKGRHRLDRERGRSPDRRODSRYSDVD 270
DB 517 PHPPEAEAR--AMIVHMDLMDIAH-----KRSRPPYRYSRSRSPPE 561
QY 271 -RGRTSPRHSYRSRERER-----ERHHRDNRSPS 302
DB 562 FRGQSTKKNVPREEKREYFNRYREVPPYDIKAYGRSVDRDPFEKERYRE----- 615
QY 303 LERSYKKEYKRSGRSYGLSVPEPACTELPGEIINKTDSWAP---PLEIVNHRSPSR 358
DB 616 WEKRYREWYKYGAVGQPRPSA-----NREDFSPERLLPLINRN--SPFT 662
QY 359 EKKRAWESEKQWSDN-----QSGKDKYTTSTKEKEPEETMPD----- 398
DB 663 RGRREDYAGQSHRNRLGNYPKLSLSTRSHAKDNPKSKESSENVDPDGKGNHKKH 722
QY 399 ----KNE---EBEBELKPVMI---RC-----THSENY--SSDPMDQGVGSTVVG 437
DB 723 RKRNRNEKEEESBSFLNPBELLETSRKCRGSSGIDETKTIDLFLVPSRDDATPVADSPMDA 782
QY 438 TS-RLDLYDKFEEL-----GSRQEKAKARPPWEPPTKLDLELSSS 481
DB 783 ESIITFSVSDKDKREKDKPVKSDKTKRSDGSATAKDNVLKPSKGPQEKVDGDEKSP 842
QY 482 ESE---CESDEDT-----CSSSDSEVFVYIAEIKKKAHPRLHDELMTYNDPGQM 530
DB 843 RSEPPPLKKAKEEATKIDSVPSSSSQDEKVTGT-----PRKAHSSKAKDTRRSQPR-- 895
QY 531 NDGPLCKSAKARTGIRHSITYPEEAIKPCRPMTNAGRLPHYRTVTSPTNFLTDRPT 590
DB 896 -----TRSRKTVPKTS-----SQKQPVTRTRPSLKIINY--LIAEKNEREKKK 941

QY 591 VIEYDDHEIYFEGSPMFAHAPLNTNPLCKVIRFNIDYTHFLEMMFENFCYKGLFSL 650
DB 942 SVDKD-----FESSSM-----KISK-----VEGTETVXP 965
QY 651 FLFR-----DILEYDMNLKGPLFEDSPCCRFHMPFRVRLPDGKEV----- 696
DB 966 SPKRMEGVDEKLETPPEKDKIASSTPA-----KKIKLNREIGKIKGNAENASTTK 1017
QY 697 -----LSMHQILLVLLR-----CSKALV----- 714
DB 1018 EPSEKLESTSKIKQEKVYKAKRKYAGSESSTLVDYTSYSGSPVRKSEKTDTR 1077
QY 715 -----PEREINMIQ-----WELDEMOKYAECK-----GMTV 742
DB 1078 RTVIRKMEBYNDNTAPAEVDVIMIQVPSKWDKDDFSEEDVDKTTQPIQSVKPSII 1137
QY 743 TNPGTKPSVRIDQDREQFNPDVITFPIIVHFGIRPAQLSYAGDPQYQKMK--SYVRL 800
DB 1138 KVVTTKPSKT-AKYTEKESEQPE-----KQKLPKEASHHELM 1173
QY 801 RHL-----ANSPVKQTDKQKLAQREBALQIRQKNTMRREVTVE-LSSQGFMT 850
DB 1174 QHELRSKSGSASSEKGRADREHSGSEKONPDKRKSGAOPDKESTVDRLSEQGHFKT 1230

Search completed: February 3, 2006, 19:43:08
Job time : 59 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:42:19 / Search time 200 Seconds
(without alignments)
2870.487 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMQGNCTCHMSFHPGRCGRGCGHAGPSPAPSPFRQNTLRLHPQOPVQYQYEPSPAS 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, AA, Main:
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	3 US-09-900-425A-2	Sequence 2, Appl1
2	7500	100.0	1374	4 US-10-079-185-2	Sequence 2, Appl1
3	7500	100.0	1374	4 US-10-774-974-2	Sequence 2, Appl1
4	7500	100.0	1374	4 US-10-805-919-2	Sequence 2, Appl1
5	7500	100.0	1374	6 US-11-001-993-2	Sequence 2, Appl1
6	7486	99.8	1374	5 US-10-723-860-281	Sequence 281, App
7	7486	99.8	1374	5 US-10-756-149-4809	Sequence 4809, Ap
8	2963	39.5	1327	6 US-11-097-143-2409	Sequence 2409, Ap
9	2847	38.0	541	4 US-10-205-331-116	Sequence 116, App
10	2463	32.8	466	4 US-10-774-974-37	Sequence 37, Appl
11	1961	26.1	378	4 US-10-103-313-471	Sequence 471, App
12	1320	17.6	263	4 US-10-103-313-307	Sequence 307, App
13	890	11.9	412	3 US-09-900-425A-3	Sequence 3, Appl1
14	890	11.9	412	4 US-10-079-185-3	Sequence 3, Appl1
15	890	11.9	412	4 US-10-774-974-3	Sequence 3, Appl1
16	890	11.9	412	4 US-10-805-919-3	Sequence 3, Appl1
17	890	11.9	412	6 US-11-001-993-3	Sequence 3, Appl1
18	340	4.5	1151	6 US-11-097-143-11586	Sequence 11586, A
19	311	4.1	560	5 US-10-450-763-51399	Sequence 51399, A
20	309.5	4.1	1557	4 US-10-437-963-137427	Sequence 137427, A
21	303.5	4.0	406	5 US-10-450-763-57609	Sequence 57609, A
22	298.5	4.0	237	4 US-10-437-963-130873	Sequence 130873, A
23	297.5	4.0	800	4 US-10-029-386-32198	Sequence 32198, A
24	292	3.9	466	4 US-10-437-963-167964	Sequence 167964, A
25	287.5	3.8	229	4 US-10-424-599-207316	Sequence 207316, A
26	285.5	3.8	258	4 US-10-437-963-159481	Sequence 159481, A
27	284.5	3.8	236	5 US-10-617-320-4394	Sequence 4394, Ap

28	284.5	3.8	1616	5 US-10-934-998-88	Sequence 88, Appl
29	282.5	3.8	445	4 US-10-437-963-161789	Sequence 161789, A
30	282	3.8	329	4 US-10-437-963-197333	Sequence 197333, A
31	281	3.7	598	4 US-10-282-122A-64954	Sequence 64954, A
32	280	3.7	266	4 US-10-425-115-189501	Sequence 189501, A
33	279.5	3.7	232	5 US-10-472-928-2508	Sequence 2508, Ap
34	278	3.7	180	4 US-10-425-115-229737	Sequence 229737, A
35	277.5	3.7	285	4 US-10-437-963-189482	Sequence 189482, A
36	277	3.7	1537	6 US-11-097-143-27759	Sequence 27759, A
37	276.5	3.7	187	4 US-10-425-115-352789	Sequence 352789, A
38	276	3.7	273	4 US-10-425-115-330708	Sequence 330708, A
39	276	3.7	948	4 US-10-308-448-15	Sequence 448, Appl
40	276	3.7	948	5 US-10-934-998-295	Sequence 295, App
41	275.5	3.7	521	4 US-10-437-963-179715	Sequence 179715, A
42	275.5	3.7	594	6 US-11-097-143-10878	Sequence 10878, A
43	274.5	3.7	399	4 US-10-437-963-124007	Sequence 124007, A
44	274	3.7	280	4 US-10-425-115-231956	Sequence 231956, A
45	273.5	3.6	431	4 US-10-437-963-204963	Sequence 204963, A

ALIGNMENTS

RESULT 1
US-09-900-425A-2
; Sequence 2, Application US/0900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-425A-2
Query Match 100.0%; Score 7500; DB 3; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMQGNCTCHMSFHPGRCGRGCGHAGPSPAPSPFRQNTLRLHPQOPVQYQYEPSPAS 60
Db 1 MMQGNCTCHMSFHPGRCGRGCGHAGPSPAPSPFRQNTLRLHPQOPVQYQYEPSPAS 60
QY 61 TTFSSNAPNPLPRPDPVPPPPMPSPAGPPLPPCPPIPPFPFNHQMRRPFPVPCPPM 120
Db 61 TTFSSNAPNPLPRPDPVPPPPMPSPAGPPLPPCPPIPPFPFNHQMRRPFPVPCPPM 120
QY 121 PPMPCPNPVPVGAAPGAGTFFPMPPSPMPHPPPVPMQOVNQQYPPGYSHNNFPP 180
Db 121 PPMPCPNPVPVGAAPGAGTFFPMPPSPMPHPPPVPMQOVNQQYPPGYSHNNFPP 180
QY 181 SFNSFONNSSFPLPSANNSSPFRHLPPYPLPKASERSPERLKGHYDDHRHDSHGR 240
Db 181 SFNSFONNSSFPLPSANNSSPFRHLPPYPLPKASERSPERLKGHYDDHRHDSHGR 240
QY 241 GGHRSGLDRRRGRSPDRRQDSRYRSDYDRGTPRRHSYERSRERERHRHRDRRS 300
Db 241 GGHRSGLDRRRGRSPDRRQDSRYRSDYDRGTPRRHSYERSRERERHRHRDRRS 300
QY 301 PSLESYKKYKRGSGYGLSVVPPAGCTPELPGELIKNTDSWAPLEIVNHRSPSRK 360
Db 301 PSLESYKKYKRGSGYGLSVVPPAGCTPELPGELIKNTDSWAPLEIVNHRSPSRK 360

```
|||||
Db 301 PSLERSYKKEYRSGSYGLSVPEPAGCTPELPGIINKNTSWAPLEIVNHSRREK 360
Qy 361 KARMBEEDKRDNDSSGDKKNYTSIKEKEPETPDKNBEEBELKPVIRCHSEN 420
Db 361 KARMBEEDKRDNDSSGDKKNYTSIKEKEPETPDKNBEEBELKPVIRCHSEN 420
Qy 421 YSSDPMQVGDSTVVGTSRLDLVDKFEELSGROEKAARPMWPKTKLDEDESS 480
Db 421 YSSDPMQVGDSTVVGTSRLDLVDKFEELSGROEKAARPMWPKTKLDEDESS 480
Qy 481 SSECEDSDSTCSSSDSEVPVIAEIRKKAHPRLHDELMYNDPGMNDGPLCKSA 540
Db 481 SSECEDSDSTCSSSDSEVPVIAEIRKKAHPRLHDELMYNDPGMNDGPLCKSA 540
Qy 541 KARRTGIRHSIYPGEBAIKPCRPMTNAGRLEPHYRTVSPPTNPLTDRPTVIEYDDHEXI 600
Db 541 KARRTGIRHSIYPGEBAIKPCRPMTNAGRLEPHYRTVSPPTNPLTDRPTVIEYDDHEXI 600
Qy 601 FEGFSMFAHAPLTNIPLCVIRFINIDYTHIFIEBMMPEMFCVKGLELFSLPFRDILEY 660
Db 601 FEGFSMFAHAPLTNIPLCVIRFINIDYTHIFIEBMMPEMFCVKGLELFSLPFRDILEY 660
Qy 661 DNMLKGPFLFEDSPCCPRFHMPFRFVRPLPDGKEVLSMHOIILYLRSKALVPEEETA 720
Db 661 DNMLKGPFLFEDSPCCPRFHMPFRFVRPLPDGKEVLSMHOIILYLRSKALVPEEETA 720
Qy 721 NMLQWELMOKYAECECKMIVTNPCTKPSVARDLDREOFNPDTVITPIYHFGIRPA 780
Db 721 NMLQWELMOKYAECECKMIVTNPCTKPSVARDLDREOFNPDTVITPIYHFGIRPA 780
Qy 781 QLSVADPOYOKLWKSYYVLRHLANSPKVKQTDOKLAOREBALOKIRQKMTREEVTV 840
Db 781 QLSVADPOYOKLWKSYYVLRHLANSPKVKQTDOKLAOREBALOKIRQKMTREEVTV 840
Qy 841 ELSSQCFMKTGIRSDVCOHAMMLPVLTNHIRYQCIMHDKLIGYFORCLQLAMTHP 900
Db 841 ELSSQCFMKTGIRSDVCOHAMMLPVLTNHIRYQCIMHDKLIGYFORCLQLAMTHP 900
Qy 901 SHHLNGMNPDBARNLSNCGIRPKYGDGRKVNMMMRKKGIINTLNINSRLGODPTPS 960
Db 901 SHHLNGMNPDBARNLSNCGIRPKYGDGRKVNMMMRKKGIINTLNINSRLGODPTPS 960
Qy 961 RINHNRELEFGDAVEFLTSVHLUYLPSSLBEGSLATYRTAIVQNHMLAKKLELDP 1020
Db 961 RINHNRELEFGDAVEFLTSVHLUYLPSSLBEGSLATYRTAIVQNHMLAKKLELDP 1020
Qy 1021 FMLYAHGPDLCRESDLRHAMANGFALLGAVULEGSLEBAKQIFGRLFNDDPLREVMIN 1080
Db 1021 FMLYAHGPDLCRESDLRHAMANGFALLGAVULEGSLEBAKQIFGRLFNDDPLREVMIN 1080
Qy 1081 YPLHPIQLOBPNDROLIETSPVLOKLEFEBAIGVIFTHVRLAFAFLRTGFGNHLTL 1140
Db 1081 YPLHPIQLOBPNDROLIETSPVLOKLEFEBAIGVIFTHVRLAFAFLRTGFGNHLTL 1140
Qy 1141 GHNQRMFPLGDSIMQVATEYLFIHPPDHHEGLTLRSGLVNNRTQAVAEELGQBEYA 1200
Db 1141 GHNQRMFPLGDSIMQVATEYLFIHPPDHHEGLTLRSGLVNNRTQAVAEELGQBEYA 1200
Qy 1201 ITNDTKRPVGLATKTLADLESFIALALYTDKOLEVYHFMNVCFFPRKEFLINDAMD 1260
Db 1201 ITNDTKRPVGLATKTLADLESFIALALYTDKOLEVYHFMNVCFFPRKEFLINDAMD 1260
Qy 1261 PMSOLOCCCLTTEGKEPDIPLYKTLQVGPSHARTYAAVFKERIGCGGSPISIOA 1320
Db 1261 PMSOLOCCCLTTEGKEPDIPLYKTLQVGPSHARTYAAVFKERIGCGGSPISIOA 1320
Qy 1321 EMGAAMDALBKYNFPMANQKRFIGRKYROELKEMMEREHOREDEDEDIKX 1374
Db 1321 EMGAAMDALBKYNFPMANQKRFIGRKYROELKEMMEREHOREDEDEDIKX 1374
```

RESULT 2

```
US-10-079-185-2
; Sequence 2, Application US/10079185
; Publication No. US20030044941A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: IS155030
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-079-185-2

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNOGNTCHMSFHPRGCGRGHGAPASFPFONRLHPOOPVYOYEPSPAS 60
Db 1 MNOGNTCHMSFHPRGCGRGHGAPASFPFONRLHPOOPVYOYEPSPAS 60
Qy 61 TTFNSPAPNLFPRPDPVFPFPPMPPSAQGLPFCPIRPPPNHMRAPFVPPCFPPM 120
Db 61 TTFNSPAPNLFPRPDPVFPFPPMPPSAQGLPFCPIRPPPNHMRAPFVPPCFPPM 120
Qy 121 PPPMPCPNPVPYGAAPGCGTFPPMPPSPMHPPPPMPOQVYQYPPGSHNFPFP 180
Db 121 PPPMPCPNPVPYGAAPGCGTFPPMPPSPMHPPPPMPOQVYQYPPGSHNFPFP 180
Qy 181 SFNSFONNSSFPLPANNSSPHFRLHPYLPKAPSEERSPERLKHVDHHRDHSOR 240
Db 181 SFNSFONNSSFPLPANNSSPHFRLHPYLPKAPSEERSPERLKHVDHHRDHSOR 240
Qy 241 GERHSLDRERGRSPDRRQDSRYRSDYDRGTFPSRHSYERSERERERHRDRNRS 300
Db 241 GERHSLDRERGRSPDRRQDSRYRSDYDRGTFPSRHSYERSERERERHRDRNRS 300
Qy 301 PSLERSYKKEYRSGRSYGLSVPEPAGCTPELPGIINKNTSWAPLEIVNHSRREK 360
Db 301 PSLERSYKKEYRSGRSYGLSVPEPAGCTPELPGIINKNTSWAPLEIVNHSRREK 360
Qy 361 KARMBEEDKRDNDSSGDKKNYTSIKEKEPETPDKNBEEBELKPVIRCHSEN 420
Db 361 KARMBEEDKRDNDSSGDKKNYTSIKEKEPETPDKNBEEBELKPVIRCHSEN 420
Qy 421 YSSDPMQVGDSTVVGTSRLDLVDKFEELSGROEKAARPMWPKTKLDEDESS 480
Db 421 YSSDPMQVGDSTVVGTSRLDLVDKFEELSGROEKAARPMWPKTKLDEDESS 480
Qy 481 SSECEDSDSTCSSSDSEVPVIAEIRKKAHPRLHDELMYNDPGMNDGPLCKSA 540
Db 481 SSECEDSDSTCSSSDSEVPVIAEIRKKAHPRLHDELMYNDPGMNDGPLCKSA 540
Qy 541 KARRTGIRHSIYPGEBAIKPCRPMTNAGRLEPHYRTVSPPTNPLTDRPTVIEYDDHEXI 600
Db 541 KARRTGIRHSIYPGEBAIKPCRPMTNAGRLEPHYRTVSPPTNPLTDRPTVIEYDDHEXI 600
Qy 601 FEGFSMFAHAPLTNIPLCVIRFINIDYTHIFIEBMMPEMFCVKGLELFSLPFRDILEY 660
Db 601 FEGFSMFAHAPLTNIPLCVIRFINIDYTHIFIEBMMPEMFCVKGLELFSLPFRDILEY 660
```

QY 661 DNMLKGPRLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRSKALVPEBEIA 720
| | | | |
Db 661 DNMLKGPRLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRSKALVPEBEIA 720
QY 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIDQDREQFNPDIYTFPIIVHFGIRPA 780
| | | | |
Db 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIDQDREQFNPDIYTFPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLKMSYVYKLRLHLANSPKVKQTDKQKLAQREBALOKIROKNTMRREVTY 840
| | | | |
Db 781 QLSYAGDPQYQKLKMSYVYKLRLHLANSPKVKQTDKQKLAQREBALOKIROKNTMRREVTY 840
QY 841 ELSSQGFMTGTGRSDVCGHAMMLPVLTTHIRYHQCIMHLDKLGTYFORCLQLAMTHP 900
| | | | |
Db 841 ELSSQGFMTGTGRSDVCGHAMMLPVLTTHIRYHQCIMHLDKLGTYFORCLQLAMTHP 900
QY 901 SHHLNFGAMPDHAARNLSNCGIRQPKYGRKYNHMMRKKGINTLININSRLGODDPTPS 960
| | | | |
Db 901 SHHLNFGAMPDHAARNLSNCGIRQPKYGRKYNHMMRKKGINTLININSRLGODDPTPS 960
QY 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGLATYRTAIVONQHLAMLAKKLELD 1020
| | | | |
Db 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGLATYRTAIVONQHLAMLAKKLELD 1020
QY 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEBAKQFGRLFNPDPLREVWLN 1080
| | | | |
Db 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEBAKQFGRLFNPDPLREVWLN 1080
QY 1081 YPLHPLQLOEPNTDROLIESPVLOKLTPEFEALIGYITFHVRLLAAPFLRTYGFNHLTL 1140
| | | | |
Db 1081 YPLHPLQLOEPNTDROLIESPVLOKLTPEFEALIGYITFHVRLLAAPFLRTYGFNHLTL 1140
QY 1141 GHNQRELEFLGDSIMQVATEYELFIHFPDHHGHLTLRSLSLVNRTQAKAYAEELGMOEYA 1200
| | | | |
Db 1141 GHNQRELEFLGDSIMQVATEYELFIHFPDHHGHLTLRSLSLVNRTQAKAYAEELGMOEYA 1200
QY 1201 ITNDTKRPVGLATKTTLADLESFIALYTDKOLEYVHTFMVCFEPRLKEPILNDQMD 1260
| | | | |
Db 1201 ITNDTKRPVGLATKTTLADLESFIALYTDKOLEYVHTFMVCFEPRLKEPILNDQMD 1260
QY 1261 PXSQLOOCCLTLTTEKEPRIPLYKTLQTYGSHARTTYAVYFKBERIGCGGSPSIQQA 1320
| | | | |
Db 1261 PXSQLOOCCLTLTTEKEPRIPLYKTLQTYGSHARTTYAVYFKBERIGCGGSPSIQQA 1320
QY 1321 BNGAAMDAL EKVPOMAHOKRIFGRKYROELKEMMEREHOREBDEDEDIKK 1374
| | | | |
Db 1321 BNGAAMDAL EKVPOMAHOKRIFGRKYROELKEMMEREHOREBDEDEDIKK 1374

RESULT 3

US-10-774-974-2

; Sequence 2, Application US/10774974

; Publication No. US20040126867A1

; GENERAL INFORMATION:

; APPLICANT: Wu, Hongjiang

; APPLICANT: Crooke, Stanley T.

; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof

; FILE REFERENCE: ISIS5029/ISPH-0522

; CURRENT APPLICATION NUMBER: US/10/774, 974

; PRIOR FILING DATE: 2004-02-09

; PRIOR APPLICATION NUMBER: US/09/900,425B

; PRIOR FILING DATE: 2001-07-06

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 2

; LENGTH: 1374

; TYPE: PR

; ORGANISM: Homo sapiens

US-10-774-974-2

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMOGNTCHNRSFHPGRCSPRGCGHAPSPAPSFRPQNTLLHPQOPVQYQYEPSAPS 60
| | | | |
Db 1 MMOGNTCHNRSFHPGRCSPRGCGHAPSPAPSFRPQNTLLHPQOPVQYQYEPSAPS 60
QY 61 TTFSSNAPNFFLPRRDVPVPPMPSPSAQGLPSPCPPIAPPFNHOMKRRHFPVPPCPPPM 120
| | | | |
Db 61 TTFSSNAPNFFLPRRDVPVPPMPSPSAQGLPSPCPPIAPPFNHOMKRRHFPVPPCPPPM 120
QY 121 PRPMPCPNPVPVGAAPGOGTFFPMMPPPSMMPHPPPPVWPOQVNTQYPPGYSHHNP 180
| | | | |
Db 121 PRPMPCPNPVPVGAAPGOGTFFPMMPPPSMMPHPPPPVWPOQVNTQYPPGYSHHNP 180
QY 181 SFNSFQNNPSSFLPSANSSSPHFRHLPPYPLPKASERERSPERLKHYPDHRHRDHSGR 240
| | | | |
Db 181 SFNSFQNNPSSFLPSANSSSPHFRHLPPYPLPKASERERSPERLKHYPDHRHRDHSGR 240
QY 241 GERHRLDRRERGRSPDRRRODSRYSDYDRGTPSGRHSYERSRERRRHHRDNR 300
| | | | |
Db 241 GERHRLDRRERGRSPDRRRODSRYSDYDRGTPSGRHSYERSRERRRHHRDNR 300
QY 301 PSLESYKKKEXKSGSYGLSVYVBPAGCTPELIGELIKNTDSWAPLEIVNHRSPSREK 360
| | | | |
Db 301 PSLESYKKKEXKSGSYGLSVYVBPAGCTPELIGELIKNTDSWAPLEIVNHRSPSREK 360
QY 361 KRARWEEKDRMSDNQSGCKKNYTSIKEKEPEETVPDKNEEBEELKPVWIRCTHSEN 420
| | | | |
Db 361 KRARWEEKDRMSDNQSGCKKNYTSIKEKEPEETVPDKNEEBEELKPVWIRCTHSEN 420
QY 421 YSSSDPMDOYGDSTVGTSLRLDLVYKFEBELGROEKAAARPWMEPPKTLDEBLESS 480
| | | | |
Db 421 YSSSDPMDOYGDSTVGTSLRLDLVYKFEBELGROEKAAARPWMEPPKTLDEBLESS 480
QY 481 SSESCESDSDSTSSSDSEVPDVIATIRKKKAPRLDELAYNPQGMNDGPLCKCSA 540
| | | | |
Db 481 SSESCESDSDSTSSSDSEVPDVIATIRKKKAPRLDELAYNPQGMNDGPLCKCSA 540
QY 541 KARTGIRHSIYGEBAIKPCRPMTNNAAGRLPHYRTVTSPTNFTLDRPTVIEYDHEYI 600
| | | | |
Db 541 KARTGIRHSIYGEBAIKPCRPMTNNAAGRLPHYRTVTSPTNFTLDRPTVIEYDHEYI 600
QY 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEBMPENPCVKGLEFSLFRDLLEYI 660
| | | | |
Db 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEBMPENPCVKGLEFSLFRDLLEYI 660
QY 661 DNMLKGPRLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRSKALVPEBEIA 720
| | | | |
Db 661 DNMLKGPRLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRSKALVPEBEIA 720
QY 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIDQDREQFNPDIYTFPIIVHFGIRPA 780
| | | | |
Db 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIDQDREQFNPDIYTFPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLKMSYVYKLRLHLANSPKVKQTDKQKLAQREBALOKIROKNTMRREVTY 840
| | | | |
Db 781 QLSYAGDPQYQKLKMSYVYKLRLHLANSPKVKQTDKQKLAQREBALOKIROKNTMRREVTY 840
QY 841 ELSSQGFMTGTGRSDVCGHAMMLPVLTTHIRYHQCIMHLDKLGTYFORCLQLAMTHP 900
| | | | |
Db 841 ELSSQGFMTGTGRSDVCGHAMMLPVLTTHIRYHQCIMHLDKLGTYFORCLQLAMTHP 900
QY 901 SHHLNFGAMPDHAARNLSNCGIRQPKYGRKYNHMMRKKGINTLININSRLGODDPTPS 960
| | | | |
Db 901 SHHLNFGAMPDHAARNLSNCGIRQPKYGRKYNHMMRKKGINTLININSRLGODDPTPS 960
QY 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGLATYRTAIVONQHLAMLAKKLELD 1020
| | | | |
Db 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGLATYRTAIVONQHLAMLAKKLELD 1020
QY 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEBAKQFGRLFNPDPLREVWLN 1080
| | | | |
Db 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEBAKQFGRLFNPDPLREVWLN 1080

QY 1081 YLHPLOLOEPNTDRLQIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRVTGFNHLTL 1140
DB 1081 YLHPLOLOEPNTDRLQIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRVTGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMOLVATEYLFIHPDHHEGLTLRSSLVNNRTQAQVAEBLQOEYA 1200
DB 1141 GHNORMEFLGDSIMOLVATEYLFIHPDHHEGLTLRSSLVNNRTQAQVAEBLQOEYA 1200
QY 1201 IINDTKRVRVGLRTKTLADLLESFIALYTDKLEVHTFMNVCFPRLKEFLINDDMD 1260
DB 1201 IINDTKRVRVGLRTKTLADLLESFIALYTDKLEVHTFMNVCFPRLKEFLINDDMD 1260
QY 1261 PXSOLQOCCLTLRTEKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCGKSPISIOQA 1320
DB 1261 PXSOLQOCCLTLRTEKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCGKSPISIOQA 1320
QY 1321 EMGAAMDALAKTNFPMQAQKRFGRKYOELKEMWEREHOREBDEDEDIKK 1374
DB 1321 EMGAAMDALAKTNFPMQAQKRFGRKYOELKEMWEREHOREBDEDEDIKK 1374

RESULT 4

US-10-805-919-2
/ Sequence 2, Application US/10805919
/ Publication No. US20040175828A1
/ GENERAL INFORMATION:
/ APPLICANT: Wu, Hongjiang
/ APPLICANT: Crooke, Stanley T.
/ TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
/ FILE REFERENCE: ISPH-0522
/ CURRENT APPLICATION NUMBER: US/10/805,919
/ PRIOR FILING DATE: 2004-03-22
/ PRIOR APPLICATION NUMBER: US/09/900,425
/ PRIOR FILING DATE: 2001-07-06
/ PRIOR APPLICATION NUMBER: US 09/479,783
/ PRIOR FILING DATE: 2000-01-07
/ PRIOR APPLICATION NUMBER: US 08/870,608
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: US 80/659,440
/ PRIOR FILING DATE: 1996-06-06
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 1374
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-805-919-2

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNOGNTCHMSFHPGCGCPRGCGHGRPSABSFRRQNLRLHLPQOPPOYOYEPPSARS 60
DB 1 MNOGNTCHMSFHPGCGCPRGCGHGRPSABSFRRQNLRLHLPQOPPOYOYEPPSARS 60
QY 61 TTFSNSPAPRFLPRRDPVFPFPPMPSAAGPLPCPIRPFPPNHRMHPFVPPCFPPM 120
DB 61 TTFSNSPAPRFLPRRDPVFPFPPMPSAAGPLPCPIRPFPPNHRMHPFVPPCFPPM 120
QY 121 PRPMPCPNPVPVGAAPGQTFPPFMMPPSPMHPPPPVPVMPQOVNYQYPPGYSHHNFPP 180
DB 121 PRPMPCPNPVPVGAAPGQTFPPFMMPPSPMHPPPPVPVMPQOVNYQYPPGYSHHNFPP 180
QY 181 SENSFONNPSGLPSANSSSPHFRLLPYPLFKASERSPPRLKHYDDHRKDSHGR 240
DB 181 SENSFONNPSGLPSANSSSPHFRLLPYPLFKASERSPPRLKHYDDHRKDSHGR 240
QY 241 GSHRSLDRERGRSPDRRQDSRYSDYRGRTPSHRSYERSRERERHHRDRNRS 300
DB 241 GSHRSLDRERGRSPDRRQDSRYSDYRGRTPSHRSYERSRERERHHRDRNRS 300
QY 301 PSLESYKXEYKSGRSGYGLSVVPEPAGCTPELPGELIKNTDSWAPPLEIVNHRSPSRK 360

DB 301 PSLESYKXEYKSGRSGYGLSVVPEPAGCTPELPGELIKNTDSWAPPLEIVNHRSPSRK 360
QY 361 KRAWEKEKRWSDQSGDKQNTYSIKEKEEETMPDKNEEBEELKPVIRCTHSEN 420
DB 361 KRAWEKEKRWSDQSGDKQNTYSIKEKEEETMPDKNEEBEELKPVIRCTHSEN 420
QY 421 YSSDPMDOVGDSTVVGTSRLDLVDKREBEISQOEKAARPMWEPPTKLDDEBESS 480
DB 421 YSSDPMDOVGDSTVVGTSRLDLVDKREBEISQOEKAARPMWEPPTKLDDEBESS 480
QY 481 SSSBESBESDSTCSSSSSEVFDVAEIKRKAHPDRDLDELWYNDPGQMDGPLECKSA 540
DB 481 SSSBESBESDSTCSSSSSEVFDVAEIKRKAHPDRDLDELWYNDPGQMDGPLECKSA 540
QY 541 KARTGIRHSIYGESEAIKPCRPMTNAGRLEFHYRITVSPPTNFTDRPTVIEYDHEXI 600
DB 541 KARTGIRHSIYGESEAIKPCRPMTNAGRLEFHYRITVSPPTNFTDRPTVIEYDHEXI 600
QY 601 FEGSFMFAHAPLTNIPLCKVIRFNIIDYTIHFTEEMMPENFCVKGLEFSLPRDILLEY 660
DB 601 FEGSFMFAHAPLTNIPLCKVIRFNIIDYTIHFTEEMMPENFCVKGLEFSLPRDILLEY 660
QY 661 DWNLKGPLFEDSPCCPRFHPFRVRLPDGKXVLSMHOILLYLRSKALVPEBEYA 720
DB 661 DWNLKGPLFEDSPCCPRFHPFRVRLPDGKXVLSMHOILLYLRSKALVPEBEYA 720
QY 721 NMLOMELEMOKYAECECKMIVTNGTKPSSYRIDOLDREQNPDPVITPPIVHFGIRPA 780
DB 721 NMLOMELEMOKYAECECKMIVTNGTKPSSYRIDOLDREQNPDPVITPPIVHFGIRPA 780
QY 781 QLSYAGDPQYOKLWMSYVLRHLANSFKVOTDOKLAQREBALOKIROKNTMREVTY 840
DB 781 QLSYAGDPQYOKLWMSYVLRHLANSFKVOTDOKLAQREBALOKIROKNTMREVTY 840
QY 841 ELSGSGFWKTGIRSVQCHAMMLPVLTHIRYHOCIMHDXIGYFODRCLLQIAMTHP 900
DB 841 ELSGSGFWKTGIRSVQCHAMMLPVLTHIRYHOCIMHDXIGYFODRCLLQIAMTHP 900
QY 901 SHLNFQMPDHAARSLNSCGIROPKYGDRAVHNMRRKKGINTLINIMSRLGQDDPTPS 960
DB 901 SHLNFQMPDHAARSLNSCGIROPKYGDRAVHNMRRKKGINTLINIMSRLGQDDPTPS 960
QY 961 RINHNERLEFLGDAVEFLTSVHLYYLPSLEEGGLATYRTAIVONOHILAMLAKKLEDP 1020
DB 961 RINHNERLEFLGDAVEFLTSVHLYYLPSLEEGGLATYRTAIVONOHILAMLAKKLEDP 1020
QY 1021 FMLYAHGPDLCESDLRHAMANCFEALIGAUVLEBSLEBAKOLFGRLFNDDPLREVMIN 1080
DB 1021 FMLYAHGPDLCESDLRHAMANCFEALIGAUVLEBSLEBAKOLFGRLFNDDPLREVMIN 1080
QY 1081 YLHPLOLOEPNTDRLQIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRVTGFNHLTL 1140
DB 1081 YLHPLOLOEPNTDRLQIETSPVLOKLTFFEEAIGVIFTHVRLARAFTLRVTGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMOLVATEYLFIHPDHHEGLTLRSSLVNNRTQAQVAEBLQOEYA 1200
DB 1141 GHNORMEFLGDSIMOLVATEYLFIHPDHHEGLTLRSSLVNNRTQAQVAEBLQOEYA 1200
QY 1201 IINDTKRVRVGLRTKTLADLLESFIALYTDKLEVHTFMNVCFPRLKEFLINDDMD 1260
DB 1201 IINDTKRVRVGLRTKTLADLLESFIALYTDKLEVHTFMNVCFPRLKEFLINDDMD 1260
QY 1261 PXSOLQOCCLTLRTEKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCGKSPISIOQA 1320
DB 1261 PXSOLQOCCLTLRTEKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCGKSPISIOQA 1320
QY 1321 EMGAAMDALAKTNFPMQAQKRFGRKYOELKEMWEREHOREBDEDEDIKK 1374
DB 1321 EMGAAMDALAKTNFPMQAQKRFGRKYOELKEMWEREHOREBDEDEDIKK 1374

RESULT 5

US-11-001-993-2
; Sequence 2, Application US/11001993
; Publication No. US20050159384A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNAse III And Compositions And Uses Thereof
; FILE REFERENCE: US1003-104 (US-5030US.D1)
; CURRENT FILING DATE: 2004-12-02
; PRIOR APPLICATION NUMBER: 10/079,185
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-001-993-2

Query Match 100.0%; Score 7500; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMQANTCHMSFPHGCGRGGHGAAPSPAPSPFQNTLLHPOOPVQYQYEPSPAPS 60
DB 1 MMQANTCHMSFPHGCGRGGHGAAPSPAPSPFQNTLLHPOOPVQYQYEPSPAPS 60
QY 61 TTFSNPAFNLPRPDPVPRPMPRPAQGLRPPCIRPFPNNHQRHPPVPRCPFP 120
DB 61 TTFSNPAFNLPRPDPVPRPMPRPAQGLRPPCIRPFPNNHQRHPPVPRCPFP 120
QY 121 PPMPCPNPVPVGAAPGCGTFFPMMPRPMRPPPPVPPVPPVPPVPPVPPVPP 180
DB 121 PPMPCPNPVPVGAAPGCGTFFPMMPRPMRPPPPVPPVPPVPPVPPVPPVPP 180
QY 181 SFNSFONNSSFPLPSANNSSPHFRLLPYPLPKAPSERSSPERLKHYYDHRHDSHGR 240
DB 181 SFNSFONNSSFPLPSANNSSPHFRLLPYPLPKAPSERSSPERLKHYYDHRHDSHGR 240
QY 241 GERHRLDRRERGRSPDRRQDSRYSDYDGRTPSRHRSYERSRERERHHRDRRS 300
DB 241 GERHRLDRRERGRSPDRRQDSRYSDYDGRTPSRHRSYERSRERERHHRDRRS 300
QY 301 PSLEERYKKEKYSKSGSYGLSVVPBPAGCTPLPGLIKNTDWAAPLEIVNHRSPSRK 360
DB 301 PSLEERYKKEKYSKSGSYGLSVVPBPAGCTPLPGLIKNTDWAAPLEIVNHRSPSRK 360
QY 361 KKAARWEEEDKRDNDSSGDKKNTYSIKKEKEPETMPDKNEEBEELIKVWIRCHSEN 420
DB 361 KKAARWEEEDKRDNDSSGDKKNTYSIKKEKEPETMPDKNEEBEELIKVWIRCHSEN 420
QY 421 YVSDMDVQDSTVVTGSRRLDYDKFEEBELSGROKAKAARPWEPKTKLDEDESS 480
DB 421 YVSDMDVQDSTVVTGSRRLDYDKFEEBELSGROKAKAARPWEPKTKLDEDESS 480
QY 481 SSBECSSDESDSTSSSSDSFVPIVIAEIKRKAHPRLHDELTYNDPGQNDPCLCKSA 540
DB 481 SSBECSSDESDSTSSSSDSFVPIVIAEIKRKAHPRLHDELTYNDPGQNDPCLCKSA 540
QY 541 KARRTIRISYIGEBAIKPCRPMTNAGRLPHYRTVSPPTPLTDRPVIYEDDHEVY 600
DB 541 KARRTIRISYIGEBAIKPCRPMTNAGRLPHYRTVSPPTPLTDRPVIYEDDHEVY 600
QY 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEEMPPENFCVGLFSLFRDILLEY 660
DB 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEEMPPENFCVGLFSLFRDILLEY 660

DB 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEEMPPENFCVGLFSLFRDILLEY 660
QY 661 DMLKGPFLFEDSPCCPRFHFMRFYRFLPDGKEVLSMHQILLYLRCSKALVPEBEIA 720
DB 661 DMLKGPFLFEDSPCCPRFHFMRFYRFLPDGKEVLSMHQILLYLRCSKALVPEBEIA 720
QY 721 NMLQWEELEWOKYAECKGMIVTNPCTKPSVIRIDQLDREQNPDVITPPIYHFGIRPA 780
DB 721 NMLQWEELEWOKYAECKGMIVTNPCTKPSVIRIDQLDREQNPDVITPPIYHFGIRPA 780
QY 781 QLSYAGDPQYOKLWKSIVKLRHLLANSPKYKQDQKLAQREBALQIKRQNTMRREVTU 840
DB 781 QLSYAGDPQYOKLWKSIVKLRHLLANSPKYKQDQKLAQREBALQIKRQNTMRREVTU 840
QY 841 ELSSQGFWMKTGISDVCQHMMPLVLTNHRHQCMLHDLKLGTFQDRCCLQAMTHP 900
DB 841 ELSSQGFWMKTGISDVCQHMMPLVLTNHRHQCMLHDLKLGTFQDRCCLQAMTHP 900
QY 901 SHHNFQAMPDHARNSLNSGIRQPKYGDRCVHMMRKKGINTLINIMSLGODDETPS 960
DB 901 SHHNFQAMPDHARNSLNSGIRQPKYGDRCVHMMRKKGINTLINIMSLGODDETPS 960
QY 961 RIMHNRLEFLGAVVEFLTSVHLVYLPSELBEGGLATYRTALVONQHLMLAKKLELD 1020
DB 961 RIMHNRLEFLGAVVEFLTSVHLVYLPSELBEGGLATYRTALVONQHLMLAKKLELD 1020
QY 1021 FMLYAHGPDLCRESDDRHAMANCFEALIGAVYLEGSLEBAKQFLFNDPDLREVMLN 1080
DB 1021 FMLYAHGPDLCRESDDRHAMANCFEALIGAVYLEGSLEBAKQFLFNDPDLREVMLN 1080
QY 1081 YPLHPLQLOEPNTDROLIETSPVLOKLTFFEEALIGVY FTHVRLLABAFTLRTVGFMHLL 1140
DB 1081 YPLHPLQLOEPNTDROLIETSPVLOKLTFFEEALIGVY FTHVRLLABAFTLRTVGFMHLL 1140
QY 1141 GHNQWMEFLGDSIMOLVATETYLFIHPDHHGHLTLRSSLVNNRTOAKVAEELGMOEYA 1200
DB 1141 GHNQWMEFLGDSIMOLVATETYLFIHPDHHGHLTLRSSLVNNRTOAKVAEELGMOEYA 1200
QY 1201 ITNDKTRPVGLTETKTLADLLESFIALALYTDKLEVYHFMNVCFFPRLKEFLINDQND 1260
DB 1201 ITNDKTRPVGLTETKTLADLLESFIALALYTDKLEVYHFMNVCFFPRLKEFLINDQND 1260
QY 1261 PKSQLQCCCLLTETGKEPDIPLKTLQTVGSHARTYTVAVYFKBERIGCGKPSIQQA 1320
DB 1261 PKSQLQCCCLLTETGKEPDIPLKTLQTVGSHARTYTVAVYFKBERIGCGKPSIQQA 1320
QY 1321 EMGAAMDALFKYFPQMAHQKRFIGRKYROELKEMWEMERHOREDEDEDIKK 1374
DB 1321 EMGAAMDALFKYFPQMAHQKRFIGRKYROELKEMWEMERHOREDEDEDIKK 1374

RESULT 6
US-10-723-860-281
; Sequence 281, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nacasha
; APPLICANT: Zlotnick, Wendy M.
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 281
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-281

Query Match 99.8%; Score 7486; DB 5; Length 1374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MMQGNVCHRMSPHPCGCRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSABS 60
DB 1 MMQGNVCHRMSPHPCGCRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSABS 60
QY 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPCIRPPPPNOMRHPFVPCFPFM 120
DB 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPCIRPPPPNOMRHPFVPCFPFM 120
QY 121 PPMPCPNPVPVGAAPGCGTFPPMMPPSMHPFPPVPMPOQVNVQYPPGYSHNFPFP 180
DB 121 PPMPCPNPVPVGAAPGCGTFPPMMPPSMHPFPPVPMPOQVNVQYPPGYSHNFPFP 180
QY 121 PPMPCPNPVPVGAAPGCGTFPPMMPPSMHPFPPVPMPOQVNVQYPPGYSHNFPFP 180
DB 121 PPMPCPNPVPVGAAPGCGTFPPMMPPSMHPFPPVPMPOQVNVQYPPGYSHNFPFP 180
QY 181 SFNSFQNNSSFLPSANSSSPHFRHLPPYPLPKASERSRSPRLKHVDHRRHDSHGR 240
DB 181 SFNSFQNNSSFLPSANSSSPHFRHLPPYPLPKASERSRSPRLKHVDHRRHDSHGR 240
QY 181 SFNSFQNNSSFLPSANSSSPHFRHLPPYPLPKASERSRSPRLKHVDHRRHDSHGR 240
DB 181 SFNSFQNNSSFLPSANSSSPHFRHLPPYPLPKASERSRSPRLKHVDHRRHDSHGR 240
QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGRTSPRHSYERSRERERHRHNRNS 300
DB 241 GERHSLDRERGRSPDRRQDSRYSDYDRGRTSPRHSYERSRERERHRHNRNS 300
QY 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPLPGELIKNTDSWAPLEIVNHRSPSRK 360
DB 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPLPGELIKNTDSWAPLEIVNHRSPSRK 360
QY 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPLPGELIKNTDSWAPLEIVNHRSPSRK 360
DB 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPLPGELIKNTDSWAPLEIVNHRSPSRK 360
QY 361 KARWEEKORWSDNOSGSKDNKNTYSIKKEPEETMPDKNEEBEELKPVIRCTHSN 420
DB 361 KARWEEKORWSDNOSGSKDNKNTYSIKKEPEETMPDKNEEBEELKPVIRCTHSN 420
QY 361 KARWEEKORWSDNOSGSKDNKNTYSIKKEPEETMPDKNEEBEELKPVIRCTHSN 420
DB 361 KARWEEKORWSDNOSGSKDNKNTYSIKKEPEETMPDKNEEBEELKPVIRCTHSN 420
QY 421 YSSOPMDVGVSTVGTSLRLDLVYKFEELGSRREKAKARPPEPKTKLDEDESS 480
DB 421 YSSOPMDVGVSTVGTSLRLDLVYKFEELGSRREKAKARPPEPKTKLDEDESS 480
QY 421 YSSOPMDVGVSTVGTSLRLDLVYKFEELGSRREKAKARPPEPKTKLDEDESS 480
DB 421 YSSOPMDVGVSTVGTSLRLDLVYKFEELGSRREKAKARPPEPKTKLDEDESS 480
QY 481 SESECESDSDSTCSSSSDESEVDVIAEIKRKAHPRLHDELYNDPGQNDGPLCKCSA 540
DB 481 SESECESDSDSTCSSSSDESEVDVIAEIKRKAHPRLHDELYNDPGQNDGPLCKCSA 540
QY 481 SESECESDSDSTCSSSSDESEVDVIAEIKRKAHPRLHDELYNDPGQNDGPLCKCSA 540
DB 481 SESECESDSDSTCSSSSDESEVDVIAEIKRKAHPRLHDELYNDPGQNDGPLCKCSA 540
QY 541 KARRTGIRHSIYGEBAIKPCRPMTNNAAGRLFHYRTVSPPTNFLTDRPTVIEYDHEYI 600
DB 541 KARRTGIRHSIYGEBAIKPCRPMTNNAAGRLFHYRTVSPPTNFLTDRPTVIEYDHEYI 600
QY 601 PEGFSMFAAPLNTITLCVIRPNIDYTHFIEEMMPENPCVKGLEFLPLFRDILEY 660
DB 601 PEGFSMFAAPLNTITLCVIRPNIDYTHFIEEMMPENPCVKGLEFLPLFRDILEY 660
QY 661 PEGFSMFAAPLNTITLCVIRPNIDYTHFIEEMMPENPCVKGLEFLPLFRDILEY 660
DB 661 PEGFSMFAAPLNTITLCVIRPNIDYTHFIEEMMPENPCVKGLEFLPLFRDILEY 660
QY 721 NMLQWELEMOQYAECKGMIVTNPCTKSSVARDLDRBOFNPDVITPIIVHGIRPA 780
DB 721 NMLQWELEMOQYAECKGMIVTNPCTKSSVARDLDRBOFNPDVITPIIVHGIRPA 780
QY 781 QLSYADPOYQKLKMSYVVLRLHLANSPKYKQDQKLAQREBALOKIKOKMTREYV 840
DB 781 QLSYADPOYQKLKMSYVVLRLHLANSPKYKQDQKLAQREBALOKIKOKMTREYV 840
QY 841 ELSSOGFMKTGIRSDVCOHANMPLVLTTHIRYQCLMLDLKLGTFODRCLLQLMTTP 900
DB 841 ELSSOGFMKTGIRSDVCOHANMPLVLTTHIRYQCLMLDLKLGTFODRCLLQLMTTP 900
QY 901 SHHLNGMNPDAHARNSLSNGIRQPKYGRKVHMMMRKKGINTLINISRLGODDPTPS 960
DB 901 SHHLNGMNPDAHARNSLSNGIRQPKYGRKVHMMMRKKGINTLINISRLGODDPTPS 960
QY 961 RINHNRLLEFLGAVVEFLTSVHLVTLFSLEREGGLATYRTALVONOHMLAKKLELP 1020
DB 961 RINHNRLLEFLGAVVEFLTSVHLVTLFSLEREGGLATYRTALVONOHMLAKKLELP 1020
```

```
QY 1021 FMLYAHGPDLCRESPLDRHMANCFEALLIGAVYLESSLEBAQLFGRLLFNDPDLREVMLN 1080
DB 1021 FMLYAHGPDLCRESPLDRHMANCFEALLIGAVYLESSLEBAQLFGRLLFNDPDLREVMLN 1080
QY 1081 YPLHFLQOEPTDQLETSPLVLOKLTPEBEAIGVIFTHVLLARAFPLRTVGFNHLTL 1140
DB 1081 YPLHFLQOEPTDQLETSPLVLOKLTPEBEAIGVIFTHVLLARAFPLRTVGFNHLTL 1140
QY 1141 GHNQMEFLGDSIMQVATEYLFIHPDHHEGHLTLSSLVNNRTQAKVAEELGMOEYA 1200
DB 1141 GHNQMEFLGDSIMQVATEYLFIHPDHHEGHLTLSSLVNNRTQAKVAEELGMOEYA 1200
QY 1201 ITNDTKRPPVGLRTTLADLBSFLAAYLTDKOLEYVNTFMNVCFPRPKETILNODMND 1260
DB 1201 ITNDTKRPPVGLRTTLADLBSFLAAYLTDKOLEYVNTFMNVCFPRPKETILNODMND 1260
QY 1261 PXSQLOCCCLTRTEGKSPDIPLYKTLQTVGSHARTVAAYFGERIGCGKSPSIQQA 1320
DB 1261 PXSQLOCCCLTRTEGKSPDIPLYKTLQTVGSHARTVAAYFGERIGCGKSPSIQQA 1320
QY 1321 EMGAAMDALIEKYNFQMAHQKRFIERKXQBELKEMWEREHQERBPDETDIKK 1374
DB 1321 EMGAAMDALIEKYNFQMAHQKRFIERKXQBELKEMWEREHQERBPDETDIKK 1374
```

```
RESULT 7
US-10-756-149-4809
; Sequence 4809, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natsasha
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patencin version 3.2
; SEQ ID NO 4809
; LENGTH: 1374
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-756-149-4809
```

Query Match 99.8%; Score 7486; DB 5; Length 1374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MMQGNVCHRMSPHPCGCRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSABS 60
DB 1 MMQGNVCHRMSPHPCGCRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSABS 60
QY 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPCIRPPPPNOMRHPFVPCFPFM 120
DB 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPCIRPPPPNOMRHPFVPCFPFM 120
QY 121 PPMPCPNPVPVGAAPGCGTFPPMMPPSMHPFPPVPMPOQVNVQYPPGYSHNFPFP 180
DB 121 PPMPCPNPVPVGAAPGCGTFPPMMPPSMHPFPPVPMPOQVNVQYPPGYSHNFPFP 180
QY 181 SFNSFQNNSSFLPSANSSSPHFRHLPPYPLPKASERSRSPRLKHVDHRRHDSHGR 240
DB 181 SFNSFQNNSSFLPSANSSSPHFRHLPPYPLPKASERSRSPRLKHVDHRRHDSHGR 240
QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGRTSPRHSYERSRERERHRHNRNS 300
DB 241 GERHSLDRERGRSPDRRQDSRYSDYDRGRTSPRHSYERSRERERHRHNRNS 300
QY 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPLPGELIKNTDSWAPLEIVNHRSPSRK 360
DB 301 PSLERSYKKEYKRSGRSYGLSVPEPAGCTPLPGELIKNTDSWAPLEIVNHRSPSRK 360
```

QY 361 KQARWEEKDRMSDNOSSGDKNYTISIKKEPEETMPDKNEEBEELKRWIRCHTSEN 420
DB 361 KQARWEEKDRMSDNOSSGDKNYTISIKKEPEETMPDKNEEBEELKRWIRCHTSEN 420
QY 421 YSSDPMDOVGVSTVGTSTRLRLYKPFEBELDSROEKAKAAPRMEPPTKLDDELSS 480
DB 421 YSSDPMDOVGVSTVGTSTRLRLYKPFEBELDSROEKAKAAPRMEPPTKLDDELSS 480
QY 481 SESECESDSDSTCSSSSSEVPFVIAEIKRKAHPRLDELYNPGQNDPPLCKCSA 540
DB 481 SESECESDSDSTCSSSSSEVPFVIAEIKRKAHPRLDELYNPGQNDPPLCKCSA 540
QY 541 KARTGTIRHSIYGEBAIKPCREPMNAGRLPFYRITVSPNTFLDRPTVEYDDHEX1 600
DB 541 KARTGTIRHSIYGEBAIKPCREPMNAGRLPFYRITVSPNTFLDRPTVEYDDHEX1 600
QY 601 PEGFSFPAHAPLNTIPLCKYIRNIDYTHFIEBMPENFCVKGELFSLFLRDLIELY 660
DB 601 PEGFSFPAHAPLNTIPLCKYIRNIDYTHFIEBMPENFCVKGELFSLFLRDLIELY 660
QY 661 DNMLKGPLFEDSPCCPRFHFMPRFVRLPDGKEVLSMHQILLYLRCSKALVPEEBE1A 720
DB 661 DNMLKGPLFEDSPCCPRFHFMPRFVRLPDGKEVLSMHQILLYLRCSKALVPEEBE1A 720
QY 721 NMLQWEELEMOXYAECECKGMITVNPCTKPSVRIIDLDREQFNPDIYTPPIVHFGIRPA 780
DB 721 NMLQWEELEMOXYAECECKGMITVNPCTKPSVRIIDLDREQFNPDIYTPPIVHFGIRPA 780
QY 781 QLSYAGDPYOYKLMKSYVTKLRHLANSPYKQOTDKOLAOREBALOKIRQKNTMBREVTY 840
DB 781 QLSYAGDPYOYKLMKSYVTKLRHLANSPYKQOTDKOLAOREBALOKIRQKNTMBREVTY 840
QY 841 ELSOSGFMKTGTIRSDVCOHAMMLPVLTHHRYHQCLMHLDKLIGYFORCLQLAMTHP 900
DB 841 ELSOSGFMKTGTIRSDVCOHAMMLPVLTHHRYHQCLMHLDKLIGYFORCLQLAMTHP 900
QY 901 SHHLNCGMNDHARNSLSNCGIRQPKYGDRCVHMHMKKKGINTLLINISRLGQDDPTPS 960
DB 901 SHHLNCGMNDHARNSLSNCGIRQPKYGDRCVHMHMKKKGINTLLINISRLGQDDPTPS 960
QY 961 RINNNERLEFLGDVAVEFLTSVHLVYLPSELEGGIATYRTAIVONQHLAMLAKKLEDP 1020
DB 961 RINNNERLEFLGDVAVEFLTSVHLVYLPSELEGGIATYRTAIVONQHLAMLAKKLEDP 1020
QY 1021 FMLYAHGPRLCRESDLRHAMANCEFALIGAVYLEGSLLEAKOLFGLLFNDPDLREVMIN 1080
DB 1021 FMLYAHGPRLCRESDLRHAMANCEFALIGAVYLEGSLLEAKOLFGLLFNDPDLREVMIN 1080
QY 1081 YPLHPLQOBNPTDRQLIETSPVLQKLTFFEEAIGVIFTHVRLAARFLLRTVGFNHLTL 1140
DB 1081 YPLHPLQOBNPTDRQLIETSPVLQKLTFFEEAIGVIFTHVRLAARFLLRTVGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMQVATEYLFIRPDHHEGHLJLLRSSLVNRTQAVAEELGQEXYA 1200
DB 1141 GHNORMEFLGDSIMQVATEYLFIRPDHHEGHLJLLRSSLVNRTQAVAEELGQEXYA 1200
QY 1201 TINDTKRVRGLRTKTLADLLESFIALYTTDKDLZVHTPMVCFEPRLKEPILNDAMD 1260
DB 1201 TINDTKRVRGLRTKTLADLLESFIALYTTDKDLZVHTPMVCFEPRLKEPILNDAMD 1260
QY 1261 PMSOLOCCULTIRTEKEPDIPLKYTLQTVGSPHATYVAVYFKERIGCGGSPSIQOA 1320
DB 1261 PMSOLOCCULTIRTEKEPDIPLKYTLQTVGSPHATYVAVYFKERIGCGGSPSIQOA 1320
QY 1321 EBGAAADALEKTNFPMAHQKRFIRGRYQELKEMWREHOREDEDEDIKK 1374
DB 1321 EBGAAADALEKTNFPMAHQKRFIRGRYQELKEMWREHOREDEDEDIKK 1374

RESULT 8
US-11-097-143-2409
; Sequence 2409, Application US/11097143
; Publication No. US20050208558A1

GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO: 2409
; LENGTH: 1327
; TYPE: PRF
; ORGANISM: DROSOPHILA
US-11-097-143-2409

Query Match 39.5%; Score 2963; DB 6; Length 1327;
Best Local Similarity 45.6%; Pred. No. 7,8e-173;
Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;

QY 90 GGPLPPCPRIPPFPNQHMHPPVPPCPMPMPMPCPN-APPVGPAP-----P 137
DB 3 QPPLPPPPVQPA-----PPPPPPPEEDLSFPQVGPVSHNYSNSNSHQ 46
QY 138 GQCTPFPMPPPSMPPPPPPVMPQOVNYQYPPGY-----SHHNPP 179
DB 47 SSKSLDYVPEPAPYASS---VPSYDPYQOPPAYGEGVAYNEQAKYGGQESHYQY 103
QY 180 PSENFQNNPSSFLPSANSSBPHRHLPPYPL---PAVSEBRREPELKYTDHRRND 235
DB 104 PA-----SGSSFLYES---YKCPDRYPAYSSNYPSPERO-----RYS 139
QY 236 HSHGGERHRSLDRRERGRSPDRRQDSRYRSDYDGRTPSHRSYERGRERERHR 295
DB 140 NSSSQGYHNP-----GYSSGR---RYEQRHQD---EHRQIQDSRYAHEPHNGY 183
QY 296 DNRSPSLERSYKKEYKSGRSGYLSVPEPAGCTPELPGELIKNTDSWAPLEIVNHS 355
DB 184 AHRQAGSQHG---YGSAAFN-----QVEDDVSPPGHHERBN 219
QY 356 PSREKKRA--RMBEEDR---WSDNOSGCKNNTSIEKEPEETMPDKNEEBEELK 409
DB 220 ETELEKTRAKPYETERDRLRLQWCSNFC-----EKREDDVYKKNALSEADAV 267
QY 410 PWMIRCTSHENYSSDPMQOVGSTVGTSTRLRLDYKPEEBELSGRQEKAKAAPRMEBP 469
DB 268 ESWNRSSPAELYTERK---SBNEYGRARQKLTLPDBELLQKAKVREGLPYVVP 323
QY 470 KTKL-----DEDLESSESSECEDESDSTCSSSDSEVPDV-----IAEIKRKAH 514
DB 324 PKARARVCXKHKKSACSSSSSSDDSDDEDA-----FKIEQDCMEELSRKVOH 373
QY 515 PDRHDELMTYNDPGQNDPPLCKCSAKARTGTIRHSIYGEBAIKPCREMTNAGRLPHY 574
DB 374 PORVHADLWHNDAGEMNDGPLCRCSAKSRIRGIRGITYGETGYKLCDPNSNNAKFLY 433

```
QY 575 RTTSPFNFLDRPVIVIEDHEEYIEGEGSMFANAPLTNIPLCIKYIRINIDYTHIFIE 634
DB 434 RISISPTNFTKTPTIKIDHEHFLBEGFSLSHVRLSDLPVCXYIRINIEVTHIEYE 493
QY 635 KMPENCVGLEFLSFLPRDILEYDMNLKGPL----BEDSPCCPRPHMFRFRFP 690
DB 494 KMPENTHIELDFFKYLHEHLELDVDMNPLBGNVEBS---CPAHFFPRFRDLP 550
QY 691 DGEKEVLSNHOILLYLRLSKALVPBEEJANMLQWEELEMOXYAECKMIVTNPQKPS 750
DB 551 DNGKEVLAVAEVLYRLDNSAQVERQOLHLNQISQSMQNVVDPIKGMALYTKPGYKPC 610
QY 751 SVRIQDLDEQRPDY-----ITPRIVHFGIRPAQLSYAGDPQYOKLMKSVYKLR 801
DB 611 STRVOLDLRN--NSDLPECVDRBTGISHPAIVHFGCHPOLSYAGNPEYOKAMREYVYK 668
QY 802 HLLANPKYQOTDKLOAREEALOKIROKNTRRREVTVELSOGQWKGISDVCOHAM 861
DB 669 HLMANMSKSPFKDKRLKEKEQLOEMRTQGRMKRITVAISSEGYRTGIMCDVYQHAM 728
QY 862 MDPVLTHNIRYHOCLMHLDKLIGYTFQDBCLQLAMTPSHHLNFGMNPDHANSLISNG 921
DB 729 LIPVLGLHLPFKSLDULEESIGYFKNRYLQLALTHPSYKENVGTNPDHANSLITNG 788
QY 922 IQPKYGDKNVHMHMKKGINTLINIMSLGODDTPSRINNELELEFGDAVVEFLTS 981
DB 789 IQPEYGDKKIHMYNTRKGINTLVSIHSRFGKEHVTNSITNELELEFGDAVVEFLTS 848
QY 982 VHLVYVPSLEBEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGPDLCRESDLBHANA 1041
DB 849 IHLFEMFELBEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGSDLCHELELHANA 908
QY 1042 NCFEALIGAVYLEGSLBEAKOLFGRLLF--NDPDLREVMNLNPLAQLQEPNTRQLIET 1100
DB 909 NCFEALMGALLDGGIKVADVEFTDALPQDEKLSITWKNLPEHPLQEBQEPIDGRCSIDS 968
QY 1101 SPVLOKLTERRBEAIGYIFTHVRLAARFILTGYFNHLLTGHNOHMEFIDGSIQOLVATE 1160
DB 969 YVLLKELTFEEDSIGIKFKHIRILARAFIDRSIGFHLTLTGSGNQRLFEPLDQLICSB 1028
QY 1161 YLFIHPRDHHEGLTLRSVNNRFOAVAEELGMOEYAL--TNDTKRPVGLRTTLTD 1219
DB 1029 YLFRHPRDHHEGLTLRSVNNRFOAVVCDLGMKPAVYANPK----ADLKTYDRAD 1084
QY 1220 LLESFIALYTDKLEEVYHFMNVCFFPRLKEFILNQDNNDPKSQLOQCCLTLLRT--EGKE 1278
DB 1085 LLEAFIAGALYVXKGLLYCQFCHVCLFPRLOLFTMNQDNNDPKSKLOQCCLTLLRTMDGE 1144
QY 1279 PDIPLYKLTQVGPSHARTYTAIVYFKGERIGCGKPSIQOEMGAMDALEKYN--FPQ 1336
DB 1145 PDIPLYKVVAEASGPTVTRVYKVAIVYFRSKRLATSSGSSIQOEMNAAKQALENSRDLFPQ 1204
QY 1337 MAHOKRFIGRKYROEL---KEMWREHOREPDEDEDIK 1374
DB 1205 LDHOKRVIAKSIKQOTGNELDNDSROHOB-----EKIKR 1239

RESULT 9
US-10-205-331-116
; Sequence 116, Application US/1020531
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alister
; APPLICANT: Brooksbank, Robert
; APPLICANT: Plimock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
```

```
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 116
; LENGTH: 541
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Putative ribonuclease III
US-10-205-331-116

Query Match          38.0%; Score 2847; DB 4; Length 541;
Best Local Similarity 99.3%; Pred. No. 3,2e-166;
Matches 537; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 834 MREYTVLSGQGFNKTGIRSDVCOHAMLPVLTNIRYHOCLMHLDKLIGTFQDBCLL 893
DB 1 MREYTVLSGQGFNKTGIRSDVCOHAMLPVLTNIRYHOCLMHLDKLIGTFQDBCLL 60
QY 894 QLAMTHPSHHLNFGMNPDHANSLISNGIROPKYGDKNVHMHMKKGINTLINIMSLG 953
DB 61 QLAMTHPSHHLNFGMNPDHANSLISNGIROPKYGDKNVHMHMKKGINTLINIMSLG 120
QY 954 ODDPTPSINNELELEFGDAVVEFLTSVHLYLFPSSLEGGIATYRTAIVONOHMLA 1013
DB 121 QDDPTPSINNELELEFGDAVVEFLTSVHLYLFPSSLEGGIATYRTAIVONOHMLA 180
QY 1014 KKLIEDPMLYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLBEAKOLFGRLLFPNDP 1073
DB 181 KKLIEDPMLYAHGPDLCRESDLRHMANCFEALIGAVYLEGSLBEAKOLFGRLLFPNDP 240
QY 1074 LREVMANPLHLOQEPNTRQLIETSPVLOKLTERRBEAIGYIFTHVRLAARFILT 1133
DB 241 LREVMANPLHLOQEPNTRQLIETSPVLOKLTERRBEAIGYIFTHVRLAARFILT 300
QY 1134 GFNHLTGHNOHMEFIDGSIQOLVATEYLFIFPDHHEGLTLRSVNNRFOAKVAEE 1193
DB 301 GFNHLTGHNOHMEFIDGSIQOLVATEYLFIFPDHHEGLTLRSVNNRFOAKVAEE 360
QY 1194 LGMQEVATINDTKRPVGLRTTLADLLESFIALYTDKLEEVYHFMNVCFFPRLKEPI 1253
DB 361 LGMQEVATINDTKRPVGLRTTLADLLESFIALYTDKLEEVYHFMNVCFFPRLKEPI 420
QY 1254 LNDNDNDPKSQLOQCCLTLLRTGKEPDIPLYKLTQVGPSHARTYTAIVYFKGERIGCGK 1313
DB 421 LNDNDNDPKSQLOQCCLTLLRTGKEPDIPLYKLTQVGPSHARTYTAIVYFKGERIGCGK 480
QY 1314 GPSIOAEMGAMDALEKYNFPQMAHOKRFIGRKYROELKEMWREHOREPDEDEDIK 1373
DB 481 GPSIOAEMGAMDALEKYNFPQMAHOKRFIGRKYROELKEMWREHOREPDEDEDIK 540
QY 1374 K 1374
DB 541 K 541

RESULT 10
US-10-774-974-37
; Sequence 37, Application US/10774974
; Publication No. US20040126867A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/10/774,974
; CURRENT FILING DATE: 2004-02-09
; PRIOR APPLICATION NUMBER: US/09/900,425B
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 466
; TYPE: PR
```

```
; ORGANISM: Homo sapiens
US-10-774-974-37

Query Match      32.8%; Score 2463; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 9.9e-143;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 909 NPDHANSLNSGIRPKYKDRKVMHMRKKGINTLINIMSRIGODDPPSRINNERL 968
DB 1 NPDHANSLNSGIRPKYKDRKVMHMRKKGINTLINIMSRIGODDPPSRINNERL 60

QY 969 EPLGDAVVEFLTSVHLYYLPSPLEEGGLATYRTAIVONQHLAMAKLEIDPMLYAGP 1028
DB 61 EPLGDAVVEFLTSVHLYYLPSPLEEGGLATYRTAIVONQHLAMAKLEIDPMLYAGP 120

QY 1029 DLCRESDLRHMANANCEBALGAVYLEGSLBEAKQLFGRLLFNDPDLREVWLANPLHPLD 1088
DB 121 DLCRESDLRHMANANCEBALGAVYLEGSLBEAKQLFGRLLFNDPDLREVWLANPLHPLD 180

QY 1089 QEPNTDROLIETSPVLOKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGNORMEF 1148
DB 181 QEPNTDROLIETSPVLOKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGNORMEF 240

QY 1149 LQDSIMQVATEYLFIFHPDHEGHLTLRSSLVNNRTQAKVAELGMOEYAITNDKTKR 1208
DB 241 LQDSIMQVATEYLFIFHPDHEGHLTLRSSLVNNRTQAKVAELGMOEYAITNDKTKR 300

QY 1209 PVGLRKTTLADLLESFIALYTDKDELYVHTFMNVCFFPRLKFIINODMNPKSQLOQC 1268
DB 301 PVGLRKTTLADLLESFIALYTDKDELYVHTFMNVCFFPRLKFIINODMNPKSQLOQC 360

QY 1269 CULTRTEGKEPDIPLKTLQTVGSPSHARTYTVAVYFKGERIGCGKPSIQOAEKGAAMD 1328
DB 361 CULTRTEGKEPDIPLKTLQTVGSPSHARTYTVAVYFKGERIGCGKPSIQOAEKGAAMD 420

QY 1329 LEKYNFPQMAHQKRFGRKYRQELKEMREREHOEREPDETEDIK 1374
DB 421 LEKYNFPQMAHQKRFGRKYRQELKEMREREHOEREPDETEDIK 466

RESULT 11
US-10-103-313-471
; Sequence 471, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 471
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-471

Query Match      26.1%; Score 1961; DB 4; Length 378;
Best Local Similarity 98.9%; Pred. No. 4.9e-112;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
DB 121 IFTHVRLARAFTLRTVGFNHLTLGNORMEFLGDSIMQVATEYLFIFHPDHEGHLTL 180
QY 1177 LRSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRTKTLADLLESFIALYTDKLEY 1236
DB 181 LRSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRTKTLADLLESFIALYTDKLEY 240

QY 1237 VHTFMNVCFFPRLKEFIINQDMNDPKSQLOQCCLTLRTGKEPDIPLYKTLQTVGSPSH 1296
DB 241 VHTFMNVCFFPRLKEFIINQDMNDPKSQLOQCCLTLRTGKEPDIPLYKTLQTVGSPSH 300

QY 1297 YTVAVYFKGERIGCGKPSIQOAEKGAAMDLEKYNFPQMAHQKRFGRKYRQELKEMR 1356
DB 301 YTVAVYFKGERIGCGKPSIQOAEKGAAMDLEKYNFPQMAHQKRFGRKYRQELKEMR 360

QY 1357 WEREHOEREPDETEDIK 1374
DB 361 WEREHOEREPDETEDIK 378

RESULT 12
US-10-103-313-307
; Sequence 307, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 307
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-307

Query Match      17.6%; Score 1320; DB 4; Length 263;
Best Local Similarity 97.7%; Pred. No. 6.7e-73;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1118 FTHVRLARAFTLRTVGFNHLTLGNORMEFLGDSIMQVATEYLFIFHPDHEGHLTL 1177
DB 7 FTHVRLARAFTLRTVGFNHLTLGNORMEFLGDSIMQVATEYLFIFHPDHEGHLTL 66

QY 1178 RSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRTKTLADLLESFIALYTDKLEY 1237
DB 67 RSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRTKTLADLLESFIALYTDKLEY 126

QY 1238 HTFMNVCFFPRLKEFIINQDMNDPKSQLOQCCLTLRTGKEPDIPLYKTLQTVGSPSH 1297
DB 127 HTFMNVCFFPRLKEFIINQDMNDPKSQLOQCCLTLRTGKEPDIPLYKTLQTVGSPSH 186

QY 1298 YTVAVYFKGERIGCGKPSIQOAEKGAAMDLEKYNFPQMAHQKRFGRKYRQELKEMR 1357
DB 187 YTVAVYFKGERIGCGKPSIQOAEKGAAMDLEKYNFPQMAHQKRFGRKYRQELKEMR 246

QY 1358 WEREHOEREPDETEDIK 1374
DB 247 WEREHOEREPDETEDIK 263

RESULT 13
US-09-900-425A-3
; Sequence 3, Application US/09900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongliang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
```

```

CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/479,783
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 08/870,608
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: US 80/659,440
PRIOR FILING DATE: 1996-06-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 412
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-900-425A-3

```

```

Query Match      11.9%; Score 890; DB 3; Length 412;
Best Local Similarity 43.1%; Pred. No. 3e-46;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;

```

```

QY 944 TLINMSRLGDDPPTSRINHNRLRLEPLGDVAVEFLTSVHLVYLPSPLEGGIATYRTAI 1003
DB 2 SLFNIMKGTSGGP-----ILHNRLRYLGDVAVELIVSHLYFMLTHHEGGIATYRTAL 57
QY 1004 VONOHIALAKKLELDPPMLYAHGPDLCRESDLRHAMANCEFALIGAVYLEGSLEBAKOL 1063
DB 58 VONRNILATLAKKCRIDEMLOQSHGADLINVAEFKALANAFBAVMAIYLDGGLAPCDVI 117
QY 1064 FGRLLF-NDPDLREVLNLYPLHPLOQEPNTDRLQLETSPLVQKLTFFEBALGVITFTHA 1122
DB 118 FSKAMYGHPVLKFKEMDHINEHELRKEDPOGDRDLSFITPTLSTFHALBERGIGQFNMR 177
QY 1123 LLARAFTLRTVGFNHLTLGHNORMEFLGDSIMQVATEYLFIHPDHEGHLTLRSSIV 1182
DB 178 LLAKAFTRRNIPNNDLTGHNORLEMLGDSVQLIVSDFLYRFPYHHEGHSILRTSLV 237
QY 1183 NNRTQAKVAEELGMOEYATINDKTRPV---GLRTKTLADLESFIALYTDKDELYVHT 1239
DB 238 SNOTQAVVCDLGFTEFVI-----KAPYKTPBELKLDKXADLVEAFIGALYVDGIEHCRA 292
QY 1240 FNNVCFPPLKEFIINQDNDPKSOLQOCCLTLR-TEGKEPDIPLXYKTLQTVGSPHARTY 1298
DB 293 FRIIVFCPLKFIIESEKNDKASHIQOQCLAMRDPSSSEPMPEYRVIGISGPTNRIIF 352
QY 1299 TVAAVFKGERIGCGSPSIQOAEWGA---AMDLEKYNFPQM-AHOKRFIGRKYROELKE 1354
DB 353 KIAVYKGRILASAEASNVAKELRVABELALANLESMSFSKMKAKNNSMRRRLBEDTSD 412

```

```

RESULT 14
US-10-079-185-3
Sequence 3, Application US/10079185
Publication No. US2003004941A1
GENERAL INFORMATION:
APPLICANT: Crooke, Stanley T.
TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
FILE REFERENCE: ISIS5030
CURRENT APPLICATION NUMBER: US/10/079,185
CURRENT FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 09/479,783
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: 08/870,608
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 08/659,440
PRIOR FILING DATE: 1996-06-06
PRIOR APPLICATION NUMBER: 09/900,425
PRIOR FILING DATE: 2001-07-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 412
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-079-185-3

```

```

Query Match      11.9%; Score 890; DB 4; Length 412;
Best Local Similarity 43.1%; Pred. No. 3e-46;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;

```

```

QY 944 TLINMSRLGDDPPTSRINHNRLRLEPLGDVAVEFLTSVHLVYLPSPLEGGIATYRTAI 1003
DB 2 SLFNIMKGTSGGP-----ILHNRLRYLGDVAVELIVSHLYFMLTHHEGGIATYRTAL 57
QY 1004 VONOHIALAKKLELDPPMLYAHGPDLCRESDLRHAMANCEFALIGAVYLEGSLEBAKOL 1063
DB 58 VONRNILATLAKKCRIDEMLOQSHGADLINVAEFKALANAFBAVMAIYLDGGLAPCDVI 117
QY 1064 FGRLLF-NDPDLREVLNLYPLHPLOQEPNTDRLQLETSPLVQKLTFFEBALGVITFTHA 1122
DB 118 FSKAMYGHPVLKFKEMDHINEHELRKEDPOGDRDLSFITPTLSTFHALBERGIGQFNMR 177
QY 1123 LLARAFTLRTVGFNHLTLGHNORMEFLGDSIMQVATEYLFIHPDHEGHLTLRSSIV 1182
DB 178 LLAKAFTRRNIPNNDLTGHNORLEMLGDSVQLIVSDFLYRFPYHHEGHSILRTSLV 237
QY 1183 NNRTQAKVAEELGMOEYATINDKTRPV---GLRTKTLADLESFIALYTDKDELYVHT 1239
DB 238 SNOTQAVVCDLGFTEFVI-----KAPYKTPBELKLDKXADLVEAFIGALYVDGIEHCRA 292
QY 1240 FNNVCFPPLKEFIINQDNDPKSOLQOCCLTLR-TEGKEPDIPLXYKTLQTVGSPHARTY 1298
DB 293 FRIIVFCPLKFIIESEKNDKASHIQOQCLAMRDPSSSEPMPEYRVIGISGPTNRIIF 352
QY 1299 TVAAVFKGERIGCGSPSIQOAEWGA---AMDLEKYNFPQM-AHOKRFIGRKYROELKE 1354
DB 353 KIAVYKGRILASAEASNVAKELRVABELALANLESMSFSKMKAKNNSMRRRLBEDTSD 412

```

```

RESULT 15
US-10-774-974-3
Sequence 3, Application US/10774974
Publication No. US20040126867A1
GENERAL INFORMATION:
APPLICANT: Wu, Hongjiang
APPLICANT: Crooke, Stanley T.
TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
FILE REFERENCE: ISIS5029/ISPH-0522
CURRENT APPLICATION NUMBER: US/10/774,974
CURRENT FILING DATE: 2004-02-09
PRIOR APPLICATION NUMBER: US/09/900,425B
PRIOR FILING DATE: 2001-07-06
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 412
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-774-974-3

```

```

Query Match      11.9%; Score 890; DB 4; Length 412;
Best Local Similarity 43.1%; Pred. No. 3e-46;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;

```

```

QY 944 TLINMSRLGDDPPTSRINHNRLRLEPLGDVAVEFLTSVHLVYLPSPLEGGIATYRTAI 1003
DB 2 SLFNIMKGTSGGP-----ILHNRLRYLGDVAVELIVSHLYFMLTHHEGGIATYRTAL 57
QY 1004 VONOHIALAKKLELDPPMLYAHGPDLCRESDLRHAMANCEFALIGAVYLEGSLEBAKOL 1063
DB 58 VONRNILATLAKKCRIDEMLOQSHGADLINVAEFKALANAFBAVMAIYLDGGLAPCDVI 117
QY 1064 FGRLLF-NDPDLREVLNLYPLHPLOQEPNTDRLQLETSPLVQKLTFFEBALGVITFTHA 1122
DB 118 FSKAMYGHPVLKFKEMDHINEHELRKEDPOGDRDLSFITPTLSTFHALBERGIGQFNMR 177
QY 1123 LLARAFTLRTVGFNHLTLGHNORMEFLGDSIMQVATEYLFIHPDHEGHLTLRSSIV 1182

```


THIS PAGE BLANK (USPTO)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:43:20 ; Search time 16 Seconds
(without alignments)
1006.288 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMQNTCHRMSPHGRGCP...MEMERHOREPDETDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276	3.7	948	US-10-523-477-14	Sequence 14, Appl1
2	260.5	3.5	245	US-10-793-626-1658	Sequence 1658, Ap
3	224	3.0	498	US-11-150-845-40	Sequence 40, Appl
4	224	3.0	498	US-11-149-945-4	Sequence 4, Appl1
5	221.5	3.0	228	US-11-074-176-136	Sequence 136, App
6	221	2.9	1560	US-11-059-982-1	Sequence 1, Appl1
7	211.5	2.8	559	US-11-150-845-38	Sequence 38, Appl
8	211.5	2.8	559	US-11-149-945-3	Sequence 3, Appl1
9	211	2.8	244	US-10-467-657-2638	Sequence 2638, Ap
10	207	2.8	795	US-10-770-726-49	Sequence 49, Appl
11	195	2.6	638	US-11-150-845-24	Sequence 24, Appl
12	194.5	2.6	501	US-11-134-563-8	Sequence 8, Appl1
13	192.5	2.6	716	US-11-150-845-16	Sequence 16, Appl
14	191	2.5	715	US-11-150-845-12	Sequence 12, Appl
15	190.5	2.5	408	US-11-150-845-12	Sequence 12, Appl
16	190.5	2.5	505	US-11-134-563-6	Sequence 6, Appl1
17	190.5	2.5	505	US-11-150-845-4	Sequence 4, Appl1
18	190.5	2.5	505	US-11-149-945-2	Sequence 2, Appl1
19	190	2.5	1823	US-10-995-561-988	Sequence 988, App
20	190	2.5	2102	US-10-995-561-990	Sequence 990, App
21	190	2.5	2108	US-10-995-561-989	Sequence 989, App
22	190	2.5	2157	US-10-995-561-991	Sequence 991, App
23	189.5	2.5	609	US-11-150-845-20	Sequence 20, Appl
24	178.5	2.4	398	US-11-150-845-10	Sequence 10, Appl
25	178.5	2.4	502	US-11-122-795-14	Sequence 14, Appl

26	178.5	2.4	502	US-11-150-845-2	Sequence 2, Appl1
27	178.5	2.4	502	US-11-149-945-1	Sequence 1, Appl1
28	178.5	2.4	649	US-11-150-845-18	Sequence 18, Appl
29	178.5	2.4	649	US-11-150-845-22	Sequence 22, Appl
30	175.5	2.3	175	US-10-821-234-1074	Sequence 1074, Ap
31	174.5	2.3	689	US-11-024-959-510	Sequence 510, App
32	174	2.3	1970	US-10-821-234-1641	Sequence 1641, Ap
33	173	2.3	502	US-11-150-845-5	Sequence 42, Appl
34	173	2.3	502	US-11-149-945-5	Sequence 5, Appl1
35	173	2.3	1532	US-10-821-234-914	Sequence 914, App
36	171	2.3	1178	US-11-044-899-29	Sequence 29, Appl
37	170.5	2.3	639	US-10-821-234-907	Sequence 907, App
38	168	2.2	580	US-10-995-561-987	Sequence 987, App
39	168	2.2	5179	US-11-108-172-1068	Sequence 1068, Ap
40	167.5	2.2	1187	US-10-821-234-955	Sequence 955, App
41	165	2.2	2804	US-11-120-925-3	Sequence 3, Appl1
42	163	2.2	2432	US-10-821-234-899	Sequence 899, App
43	162	2.2	572	US-10-821-234-1230	Sequence 1290, Ap
44	162	2.2	663	US-11-024-959-314	Sequence 314, App
45	157.5	2.1	363	US-11-054-281-120	Sequence 120, App

ALIGNMENTS

RESULT 1
US-10-523-477-14
; Sequence 14, Application US/10523477
; Publication No. US20050266406A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MAXS AS MODIFIERS OF THE AXIN PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-051C-US
; CURRENT APPLICATION NUMBER: US/10/523,477
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,534
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,153
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 948
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-523-477-14

Query Match 3.7%; Score 276; DB 6; Length 948;

Best Local Similarity 23.0%; Pred. No. 4e-10;
Matches 149; Conservative 61; Mismatches 208; Indels 230; Gaps 31;

QY	27	ARPSA--PSFRQNLRLHPQCPVQYQYEPBGA-----PSTFNSPAPNPLPR	75
DB	320	ARRGGRPQWEMSHN-FLGVLSPBQPIRGERSCYNSINGRHHSRSQRTQGS-LPNT	377
QY	76	PDVFPVPPMPPSAQGLPCCPIRPPFNHOMKHPVPVPCFPMPBPMPCPNPPVPGA	135
DB	378	PVFPVFPVPP-----PLXPFP-----HTLPPLP-----PGVPPQGSPPFP--PQ	415
QY	136	PPQCGTFPPMPMPSPNHPPPVMPQVYN-----YQYPP-----	170
DB	416	PPDAG---YSVPDPGP--DAPNVLSTPWSGQVTAHNTIPTQAPLRSREFFREQR	471
QY	171	-----	184
DB	472	LKEEKKKSLDEFTNDFAKELMEYKIKERRRSRSKSPYSGSYRSSTYSKSN	531
QY	185	PQNPSSFLPSANSSSPHFRLLPPLPKAPSERSPRLKHYDDHRRDRSHGRGERH	244
DB	532	GSTRKRSYRSRSRSHRSYRSRSPVY-----RRGRKSNY-----RRSRSHG---YH	578
QY	245	RSIDRRERGRSPDRRQDSRYNSDYD-RGRTPSRHRSYRSRERER-----	289

```

Db      579 RS-----RSRSPPPRRYHSRSRSPQAFRCQSPNKRNVPOGETEREYFNRYREVPPEYDMK 633
Qy      290 -----ERRHRDRNRSPSLERSYKKGSRSGYGLSVAPPACTBEPGE 336
Db      634 AAYGASVDRDPPEKRYE-----WEKRYEWYKYYKGAAGQPPPSA----- 679
Qy      337 IIKNTDSMAP-----PLEIVNHSRSPSEKRRARWEEKRWSNDQSG-----KD 381
Db      680 ---NRNFSRPERPLINIRN--SPFTRGRREDVYGQSRHSRNICGNYDEKLSARDGHNQ 734
Qy      382 KNYTSLKEKEPEBTMD-----KNEEERELKPPWIRCTHSENYSSDPM-- 427
Db      735 KONTSKSEKESAPDQGNKKRKRKRKGESESGFLNPELLETSRK---SRPITGV 790
Qy      428 -DQVGDSTVGTSR-----LRDYDKFEE---LGSROKAAKARPPWEPPTKLDLEDLE 478
Db      791 EENKTDLSLVPSRDATVPRD--EPMDAESTIPKSVSEKDKRER---DKPRAKGDKTKR 845
Qy      479 SSSESECEDESDSTCSSSDSE-----VPDVIAIRKKAHPDRL 518
Db      846 KNDGSASVSKENIVKPAKGPOEKVDGDVRLDLNLQLKKPKRRLRL 893

```

RESULT 2

```

US-10-793-626-1658
; Sequence 1658, Application US/107931626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMELLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1658
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1658

```

Query Match 3.5%; Score 260.5; DB 6; Length 245;

Best Local Similarity 30.4%; Pred. No. 7.2e-10; Matches 70; Conservative 57; Mismatches 88; Indels 15; Gaps 8;

```

Qy      1105 OKLTFEEAIGVIFTVRLRLARAFTLRV--GFNHLTLGHNMREFLGDSIMQLVATEYL 1162
Db      15 OKFTDMKSLGLRFXKIDLYQCAFSSHSPINDFNMRLENELBELGLDAVELYTSRL 74
Qy      1163 FTHFPDHHEGHTLLRSVLNNFTQAKVAELGMOEYATINDTKRPVGLRTK--TLADL 1220
Db      75 FDRHPLPGLNLTKMAATVTCESPLVTFANKIKLNLILLG--KGEKGTGRTTPSLISDA 133
Qy      1221 LBSFIALYTDKDLKLEVHFPMVNCPPRLK--EFLINQMNDRKSLQLOCCCLTRTEGEE 1278
Db      134 FEFVALYLDQGLDSVMTFAEKVTFPPYEDDLVGVVPF--KTFQF--YVHSQNK- 186
Qy      1279 PDIPLKTLQTVGPSHARTYTVAVVFKGERIGCGKSPSIQOAMGAMDA 1328
Db      187 GDV-TYQLKEBSPAHRLFTSEVILBNKAVAGKKTKESEKXAEOA 235

```

RESULT 3

```

US-11-150-845-40
; Sequence 40, Application US/11150845
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokineics, Inc.

```

```

; APPLICANT: Tomasevici, Nenad
; APPLICANT: Uia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-007720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 40
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(498)
; OTHER INFORMATION: SCAR2/WAVE2
US-11-150-845-40

```

Query Match 3.0%; Score 224; DB 7; Length 498;

Best Local Similarity 29.5%; Pred. No. 3.7e-07; Matches 76; Conservative 17; Mismatches 79; Indels 86; Gaps 16;

```

Qy      19 PRGRGHARAPAPSPFRQNLRLHPOQPPVGYQ-----EPP 56
Db      248 PPGQSDASSPS--PFSSEDLN-----PPPAFSTPVNDNQSGLAGPGRSSVVSBSHP 301
Qy      57 SAPSTFNSPAPNLPRLPRDPFVPPPPMPSPSAQG--PLRCPD-----RPPFPHQMKHPP 112
Db      302 PAPPLGSPGPKGAP-----PAPPPPPMIGIPPPPPVGGSPGTP-----PP 350
Qy      113 VPPCFPPMP---PPMPCNNPVPQAPPGQCTFPPMAPP--SMP---HPPPPVMPQOV 164
Db      351 SPPSFPHPDFAAPP---PPPPA-----DYPTLPPPLSGPTGAPPPPPPP 398
Qy      165 NYQVPRGYSHNFPSPSPFSPNMPSSFLPSANSSPHFRLLPYPLPKASER----- 219
Db      399 ---PPGP-----PPPTGADGQPAIPPLSDTTK-----PSSLPVADSANDSLLS 442
Qy      220 --RSPERLKHYYDHRHRD 235
Db      443 ATRQGFQLARVEGREGQE 460

```

RESULT 4

```

US-11-149-945-4
; Sequence 4, Application US/11149945
; Publication No. US20060014266A1
; GENERAL INFORMATION:
; APPLICANT: Cytokineics, Inc.
; APPLICANT: Tomasevici, Nenad
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: PURIFICATION OF ARP2/3 COMPLEX AND COMPOSITIONS CONTAINING
; FILE REFERENCE: 020552-007810US
; CURRENT APPLICATION NUMBER: US/11/149,945
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/578,969
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 4
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-149-945-4

```

Query Match	3.0%	Score 224	DB 7	Length 498
Best Local Similarity	29.5%	Pred. No. 3.7e-07		
Matches	76	Conservative 17	Mismatches 79	Indels 86
			Gaps 16	
Qy	19	PRGGGCGARSAASFRPQNLRLHPRQRPVQYQY-----	---EPP	56
Db	248	PPPOSDSASSTG--DSFSEDLN-----PPPPAEFSTYVDNQSGSLAGPRKSSVSPSHPP		301
Qy	57	SAPSTTSSNSPAPNPLPRPDPFVFRPMPSPSAQG--PLRPCH---RRFPNHQMRHPP		112
Db	302	PAPPLSGSPGCKPGKAPAP-----PAPPPRRPMIGIPRRPPVGGSGGTP-----PPR		350
Qy	113	VPSPCFPMPP---DPMPCNNRPVPGADPGGTFPMMRP--SMP--HPPRPVMPQOV		164
Db	351	SPSPFPHPHDPALAPPP---PPPPAA-----DYPTLPPPLPGTGGAPPPPPPP---		398
Qy	165	NYQVPRGYSHHNPFRPSPFNSFQNNPSSFLPSANNSSSHFRLPPLPYLPKASER-----		219
Db	399	---PPGG---PPPPFTGADGQPAIPPLPSDTTK-----PKSSLPAAVSADASDLSL		442
Qy	220	--RSRERLKHVDDHHRD	235	
Db	443	AIRQGFQLRVVEQREDE	460	

```

RESULT 5
US-11-074-176-136
; Sequence 136, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAlliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 228
; TYPE: PR1
; ORGANISM: Lactobacillus acidophilus
; US-11-074-176-136

```

	Query Match	3.0%; Score 221.5; DB 7; Length 228;
	Best Local Similarity	29.0%; Pred. No. 26-07;
	Matches 67; Conservative	47; Mismatches 102; Indels 15; Gaps 6
Qy	1107 LTFEEBAIGVIFTRVRLTLAFTLRATGFPNHLTG--HNQMEFLGDSIMOLVALEYLFI	1166
Dz	7 LNKKNENYIEFNNEKLEAEAFTHSSYSNEHPDDGIRDYEKLFEGDAVLTAIVSNLYLR	66
Qy	1165 HFRPHHEGHILFSLVNRRTOAKVAELDMQ--EYALTNRKTRPVGLRRTTLADLES	1223
Dz	67 HYPLDNGEGLTRLRNSIVRTGEFSFALECQGEIHLDGNEBEKGARNRKLTLEDVEEA	126
Qy	1224 FIALLVNDKDLEVYHTEFMNVCFPPRLK--EFFIINDMNMDPKSQTQQCCILTETSEKEPDI	1281
Dz	127 FNGALFLDDCGMPAVERFLHLTVLPIDANGEF---DDSDRYKTDLDE----LLQGNGPANI	179
Qy	1282 PLYTTLDTGVESHARTYVAVYFKGERICGGKPSIQGAENGAMDALREKY	1332
	180 EYSAVISSTOLFSSH---FVELKLINDQVOVGEGNHKKAAEQQAALAQKF	227

RESULT 6
US-11-059-982-1

```

? Sequence 1, Application US/11059982
? Publication No. US2005025507A1
? GENERAL INFORMATION:
? APPLICANT: Jenkins, Robert B.
? APPLICANT: Yang, Ping
? APPLICANT: Thibodeau, Steve
? APPLICANT: Wang, Liang
? APPLICANT: Schaid, Daniel
? TITLE OF INVENTION: CYTOGENETICALLY DETERMINED DIAGNOSIS AND
? TITLE OF INVENTION: PROGNOSIS OF PROLIFERATIVE DISORDERS
? FILE REFERENCE: 07039-505001
? CURRENT APPLICATION NUMBER: US/11/059,982
? CURRENT FILING DATE: 2005-02-17
? PRIOR APPLICATION NUMBER: US 60/545,573
? PRIOR FILING DATE: 2004-02-17
? NUMBER OF SEQ ID NOS: 50
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 1560
? TYPE: PRT
? ORGANISM: Homo sapiens
US-11-059-982-1

Query Match          2.9%   Score 221;   DB 7;   Length 1560;
Best Local Similarity 21.7%   Pred. No. 2,4e-06;
Matches 166; Conservative 55; Mismatches 231; Indels 314; Gaps 37

```

Query Match	2.9%: Score 221; DB 7; Length 1560;
Best Local Similarity	21.7%: Pred. No. 2.4e-06;
Matches 166; Conservative	55; Mismatches 231; Indels 314; Gaps 37;
Qy	4 GNTCHRSFPHGRCSPRGCGHGAPASPSFRPONTLLHNPQDPVOYQYBPSPST- 61
Db	786 GDSPHLPSPHPTR-----PSPRPSPQSVS-RPSPBPL-HPCBPQAPPLP 832
Qy	62 ---TFNS-----PAPNPLPRDPFVPPPPMPSAQ---GPLRCPPLRPFFNH----- 105
Db	833 GIFVIOQLQGVPPASNPAPTPAG--PPQPLRQSQOPBEPPLPAPLPPSSSTSAVAS 890
Qy	106 ---QMBPFPVPFCPPMPMPMPCNNPVPVAPAPGQGTFFPMPPSMHPPPP----- 157
Db	891 SSETSSLLPAPTPSDFOQLPPSPGPHKSPRP--PPLT-----HLVPEAAAPPPRRPFRQ 944
Qy	158 ---EVMPOQ-----PPLT-----VNYQ 167
Db	945 MVTTPPALPQPKALTERFHQVPSGILLQNKAGAPAPQTSSTLGPLTSPASVLSGQ 1004
Qy	168 YPPPY---SHHNPP-----PSPNSPQNNSSPLPSAN-----NSSPHPHLPYP 211
Db	1005 APSTGTPAPSAAPAPMAATGTPPLLPALNKAPASMLPTLVAKAASSGGKSGCLQYE 1064
Qy	212 ---LPAESERSRPER-LKHYDDHR---HRDHSHG---RGERHRSIDRR-ERGRS 255
Db	1065 SKLSGLKKPPLLOSKKACFLYHLHKQGSVLHNDYKTAPRPSFDALHRLLPYHYGAL 1124
Qy	256 PDRRRQDSRYSDYDGRTPBSRHSYRSRERERERHNR-----DNRR-SPSLF----- 304
Db	1125 PS-----PSDYHKYDEBEETVSTOLKTKTQMLMKYRLLLEESRRVSPSAEMWID 1176
Qy	305 ---RSYKKEYKSGSYGSVLPVDEPAGCTPRLPGELIKNTDWSAP 346
Db	1177 RMFLQEBKTTLLADKOLAKEKPEDEYVSSNSLGLPIAASSG--HRLPGH--GLSSSAP 1232
Qy	347 ---PLBYNHRSPSRREKKRARWEBEKRDWSDNOSGDKXNYTSIKEKEPEETM 396
Db	1233 GASTQPPPHLPYKLVINHGGAGSPVTWPAASSLSSSSSSAASLADDEDGP---M 1289
Qy	397 PDKKE-----EEBELLKPVW----- 412
Db	1290 PSRRRPEIKTYEARSRIQLKIKIOEAGLSKVYNTALDPVHOPPRPATLKVAPRRPR 1349
Qy	413 ---IRCHSENYYSDMDQVG----- 431
Db	1350 PPPPPPGQMGNGVYDHPPPAPAPKPKGTAPAHCRRLFLKTYREN-----VGGGGA 1400
Qy	432 -DSTVVGTSR-----LRDLYDKFEELGSRQEK-----AKAAR- 463

Db 1401 PEGTPAGRAGGSPAPLPKVDATSGILIREL-AAVEDELYORMLKGPPEPAASAAGCT 1459
Qy 464 --PMEPPKTKLDEDESSSESCSEDESDSTCSSSSSESVFVIAE 507
Db 1460 GPDWEAP-----GLPPAKRRKSESPPVDQASFSFSDSQDDTLTE 1499

RESULT 7

US-11-150-845-38
Sequence 38, Application US/11150845
Publication No. US20060003399A1
GENERAL INFORMATION:
APPLICANT: Cytokineics, Inc.
APPLICANT: Tomasevic, Nenad
APPLICANT: Jia, Zhiheng
APPLICANT: Sakowicz, Roman
APPLICANT: Pierce, Daniel
APPLICANT: Finer, Jeffrey
TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
FILE REFERENCE: 020552-007200S
CURRENT APPLICATION NUMBER: US/11/150,845
CURRENT FILING DATE: 2005-06-10
PRIOR APPLICATION NUMBER: US 60/673,444
PRIOR FILING DATE: 2005-04-20
PRIOR APPLICATION NUMBER: 60/578,949
PRIOR FILING DATE: 2004-06-10
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.3
SEQ ID NO 38
LENGTH: 559
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(559)
OTHER INFORMATION: SCAR1/WAVE1
US-11-150-845-38

Query Match 2.8%; Score 211.5; DB 7; Length 559;
Best Local Similarity 25.7%; Pred. No. 2.7e-06;
Matches 86; Conservative 30; Mismatches 111; Indels 107; Gaps 16;

Qy 35 RPONTRLHPQOPVQYQYEP-----SAPSTTFNSNP-----APNFL 72
Db 266 RAERVLVVRHPEP-----PPPMHGAGDAKPIPTCISATGLIENRPQSPATGRFPV 320
Qy 73 PRPDPVPPPPMPPSA-----QGILPCCIRPPPHQMRHPVPVPCFP--- 118
Db 321 SPTP---PPPPPLPSALSTSLRASMTSTPPPPVPPPPPPPTALQAPAVPPPAILOI 377
Qy 119 -----PMPPMPCP---NNPVGAPPGQGTFFMMPP-----SMHPPPPVMPQOV 164
Db 378 AEGVLHAPPPPIAPLVQSPAPARAAPVCETVP-VHPLQSGVOGLPPPPPPPLP--- 433
Qy 165 NYQPPGYSHNPPPSFQNNPSSFLPSANSSSPHFRLLPYVP---LPKAPSERR 220
Db 434 ----PPGIRPSS--PVTVALAHPPSGLHPTSTAGPHVPLMPSPSPQVTP-ASEPKR 486
Qy 221 SPERLKHVDHHRDHSHGGRHRSGLDRERGRSPDRRQDSRYSDYDRGTPSRHRS 280
Db 487 HSTLPLVISD-----ARSVLEAIRKG-----IQLRK 513
Qy 281 YERSERERERHRHNRSPSLERSYKKEYKRS 314
Db 514 VEEQREOE-AKHRIENDVATILSRRIAYEYSDS 546

RESULT 8

US-11-149-945-3
Sequence 3, Application US/11149945
Publication No. US20060014266A1
GENERAL INFORMATION:
APPLICANT: Cytokineics, Inc.

APPLICANT: Tomasevic, Nenad
APPLICANT: Sakowicz, Roman
TITLE OF INVENTION: PURIFICATION OF ARP2/3 COMPLEX AND COMPOSITIONS CONTAINING
TITLE OF INVENTION: PURIFIED ARP2/3 COMPLEX
FILE REFERENCE: 020552-0078100S
CURRENT APPLICATION NUMBER: US/11/149,945
CURRENT FILING DATE: 2005-06-10
PRIOR APPLICATION NUMBER: US 60/578,969
PRIOR FILING DATE: 2004-06-10
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.3
SEQ ID NO 3
LENGTH: 559
TYPE: PRF
ORGANISM: Homo sapiens
US-11-149-945-3

Query Match 2.8%; Score 211.5; DB 7; Length 559;
Best Local Similarity 25.7%; Pred. No. 2.7e-06;
Matches 86; Conservative 30; Mismatches 111; Indels 107; Gaps 16;

Qy 35 RPONTRLHPQOPVQYQYEP-----SAPSTTFNSNP-----APNFL 72
Db 266 RAERVLVVRHPEP-----PPPMHGAGDAKPIPTCISATGLIENRPQSPATGRFPV 320
Qy 73 PRPDPVPPPPMPPSA-----QGILPCCIRPPPHQMRHPVPVPCFP--- 118
Db 321 SPTP---PPPPPLPSALSTSLRASMTSTPPPPVPPPPPPPTALQAPAVPPPAILOI 377
Qy 119 -----PMPPMPCP---NNPVGAPPGQGTFFMMPP-----SMHPPPPVMPQOV 164
Db 378 AEGVLHAPPPPIAPLVQSPAPARAAPVCETVP-VHPLQSGVOGLPPPPPPPLP--- 433
Qy 165 NYQPPGYSHNPPPSFQNNPSSFLPSANSSSPHFRLLPYVP---LPKAPSERR 220
Db 434 ----PPGIRPSS--PVTVALAHPPSGLHPTSTAGPHVPLMPSPSPQVTP-ASEPKR 486
Qy 221 SPERLKHVDHHRDHSHGGRHRSGLDRERGRSPDRRQDSRYSDYDRGTPSRHRS 280
Db 487 HSTLPLVISD-----ARSVLEAIRKG-----IQLRK 513
Qy 281 YERSERERERHRHNRSPSLERSYKKEYKRS 314
Db 514 VEEQREOE-AKHRIENDVATILSRRIAYEYSDS 546

RESULT 9

US-10-467-657-2638
Sequence 2638, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASTIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMin99, version 1.04
SEQ ID NO 2638
LENGTH: 244
TYPE: PRF
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2638

Query Match 2.8%; Score 211; DB 6; Length 244;
Best Local Similarity 28.3%; Pred. No. 1e-06;
Matches 68; Conservative 41; Mismatches 117; Indels 14; Gaps 6;

```

QY      1095  ROLLETSVVL--OKLTREEEAAIGVIFTRVRLRAAFRLRTVGFPHNLTLGHNOBFLGDS 1152
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2      RKTVMKQDVLGRQAHNTALQKQLGYAFRDMSLRLRLTLTRS--HNAK--HNERPEFVDS 56

QY      1153  IMQVATEYLFIFHPDHHGHLLTLRSLVNNRTOAKVAEELGMQEVATITDKTRPVL 1212
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      57      ILNTVARMFLFDAFPLKTEGELSRLRAJLVNEGVLAEWAEMVNDGGLYLGAELKSGF 116

QY      1213  RTKI-LADLESFIALYTDKQLEVTFFMNVCFPRFLKEPLFLNDNMNDPSQLQOCLT 1271
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      117      RRPSTLADAMEAMFAAASFPDADFNTAEKVRHLLFERVRADFPONQAKDKTALQEA--- 173

QY      1272  LRTGKEBPDIPLYKTLQTVGSPHARTYTVAVYFKERIGC-GKGPSIOAEGAMADALE 1330
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      174  --LQARRALKRYIEQIGHADSMFYISCDLGBLGFVCAKGTSRRAAEQBAKELK 231

RESULT 10
US-10-770-726-49
Sequence 49, Application US/10770726
Publication No. US20050266409A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Brown, Eugene
APPLICANT: Liu, Wei
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
TITLE OF INVENTION: CANCERS
FILE REFERENCE: AM101079 (031896-010000)
CURRENT FILING DATE: 2004-02-04
NUMBER OF SEQ ID NOS: 48640
SOFTWARE: PatentIn version 3.2
SEQ ID NO 49
LENGTH: 795
TYPE: PRT
ORGANISM: Homo sapiens
US-10-770-726-49

Query March 2, 8%; Score 207; DB 6; Length 795;
Blast local similarity 26.4%, Pred. No. 8e-06;
Matches 89; Conservative 41; Mismatches 127; Indels 80; Gaps 13;

QY      214  KAPERSRSPEFLKHYDD-----HHRDHS-----HGRGBRRH-SLDRRE 251
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      86  KPPQOMSKKEVYHHRKDKKEKRRKRRSHSABGGHAAVKEKEREHERKXKHREODYAR 145

QY      252  RGRSPDRRRQDSRYRSDYDRGRTSPRHSYERSR-----ERER--ERHRRH 295
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      146  REMERQKRREWARHESRRERDLQLERKRERERMRQEQKEQREPKERRRERARRKER 205

QY      296  DNRGSPSL-EKSYKKEYKSGRSYGLSVPRPAGCTPLPBEIITKNTDSMAPPLEIVNHR 354
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      206  EAAREVAAHHTHMEYSDKVKASHWSRSP-----PPPRERPELGDGRP-----GEA 254

QY      355  SPSEKKRARRKEERK---DKMSDNQSGKQKNTYSIEKEPPEFTMPDNKEEBEELIKP 410
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      255  RPAADQKPAQJKEKMEERDLSDLOQDISDERKTSNAESSAASGSSSEEBEBS----- 310

QY      411  VMIRCHSENYSYSDPMDQVDSIVVGTSTRLRLDYLKEEELGSRQEKAKAAPRWEPBK 470
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      311  -----EEEBEEG-STSESESEEBEEBEEBEETGVSBEA-SEQSAEEVSE 354

QY      471  TKLDEDLSESES-----ECEDEDSTCSSSSDSSEV 501
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      355  EEMSEDEBERENENHLLVVPESRFPDRDSGESSSEAEV 391

RESULT 11
US-11-150-845-24
Sequence 24, Application US/11150845
Publication No. US2006000399A1
GENERAL INFORMATION:

```

```

; APPLICANT: Cytokinetics, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Flier, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-007720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MYC-98N-WASP-TAP
US-11-150-845-24

Query Match 2.6%; Score 195; DB 7; Length 638;
Best Local Similarity 23.3%; Pred. No. 3.5e-05;
Matches 102; Conservative 40; Mismatches 177; Indels 118; Gaps 21

QY 52 QYEPSPASPTTFSSNAPNFLPPRPDVPFPMPSPASQGLPCEPIK-----PPFPNH 105
DB 202 QAPPPPPS---RGPPPP---PPPPHNSGPPPPPPAGRGAPPPPPSAPTPAAPPPPPS 255
QY 106 QMRHPFPVPP---CFPPMPMPMP---CPNNPVPGAPPG---QCTFPFMPMPSPMHPPPPV 159
DB 256 RPSVAVPPPPPPNMYPPPPPALPSSASGPPPP---PSSVLCGVAVAPPP---PPPPPPPG 311
QY 160 MPQOVNVQYPPGY---SHNPPPPSPFSSF---QNNSSFLPSANNSSP----- 202
DB 312 PPP-----PPGLPPSDGDHVPPTTAGKAKALLDIREGAQQLKKVYQNSRPVSCSGRDALL 365
QY 203 -HFRH-----LPPYPLPKAPSERSPERLKHVDHRRHDSHGRGRRSLD 248
DB 366 DQIRQGIQLSVADGGQSTPTPTPAFTSGIYGALMEWQK-----RSKAIHSD 413
QY 249 RREKGRSPDRRRODSRYSRD---YDRGTPPSRRHSYERSRERERHRRHNRSPSLE 304
DB 414 EDEBDEDEBDFEDDDEWEDDPAFLYKVVDIPSHWRPLE-----TMEKRWKKNFLAVSAA 468
QY 305 RSYKGYKRSGRSYGLSVBEPACSTDELPGELIKNTDSWAPLEIYNHRSPSREKKRAR 364
DB 469 NREFKISSSGALDYD---IPTTASENLVFOGEL---KTPAALLAQHDEAVDNKP---NKEQONAF 522
QY 365 WE-----EKDMSDNQSSGCKQNYISIKEKEEE 394
DB 523 YEILHLPLNLEBQGNAPFIOQLKDDPSQSANLLAEAKKLDNQAPKVDNKFNKEQONAFYE 582
QY 395 T--MPDNKEESEEELK 409
DB 583 IHLPLNLEBQGNAPFIQ 599

RESULT 12
US-11-134-563-8
; Sequence 8, Application US/1134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campellone, Kenneth G.
; TITLE OF INVENTION: ESPFV NUCLEIC ACIDS AND PROTEINS AND
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20

```


QY 272 GTPSRHRSYERSRER--ERERHHRDRSRSPSLERSYKKEYKSGRGYGLSVDEPAG 328
DB 510 DQAGDEDEDEDEWDEPLETEMEKRWKGFIVASANRPKTISSGALDYD--IPTAS 566
QY 329 CTPELBGEIITKTDSWAPLEIVNHSRPREKKRARE----- 366
DB 567 ENLYFOGEL--KTAALAQHDEAVDNKF--NKEQONAFYEILHLPNLNEQORNAFIQSLKDD 623
QY 367 -----BEKDRWSDNQGSKDKNYTSIKEKEPEET--MPDKNEEEELLK 409
DB 624 PQSANLALAEAKKLNDAQAPKYDNKFNKQONAFYEILHLPNLNEQORNAFIQ 676

RESULT 15

US-11-150-845-12
; Sequence 12, Application US/11150845
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokine, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-00720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 408
; TYPE: PRP
; ORGANISM: Homo sapiens
US-11-150-845-12

Query Match 2.5%; Score 190.5; DB 7; Length 408;
Best Local Similarity 34.4%; Pred. No. 3.9e-05;
Matches 52; Conservative 9; Mismatches 59; Indels 31; Gaps 7;

QY 19 PRGRGSHGARPSAPSRPQNLRLHPQRPVQYQYERPSAPSTTFNSNPAAPNLPRRDPF 78
DB 184 PPSRGG---PPPPPPPHN---SGPPPPARGRGAPPPPSAPTAAPPP--PPSRPS 234
QY 79 VFPPPPMPSAQGLRPPCFIRPPFPNHNQMRHPPRVPPCFPPMPPMPCNNPPVPGAPPG 138
DB 235 VAVPPP-PNNKMYPPPP---PALPSS-----APSGPPPPPSVLGVGVAAPPPPP- 280
QY 139 OGTFPPMPPSMHPPPPPPVMPQOVNYQY 169
DB 261 -----PPPPPPGPPPPGGLPSSDGDHQP 303

Search completed: February 3, 2006, 19:46:56
Job time : 17 secs

THIS PAGE BLANK (USPTO)